DOCUMENT RESUME

ED 166 962 FL 010 068

AUTHOR	Oskarsson, Mats
TITLE	Approaches to Self-Assessment in Foreign Language Learning.
INSTITUTION	Council for Cultural Cooperation, Strasbourg (France).
PUB DATE	78
NOTE	61p.
AVAILABLE FROM	Council of Europe, F-67006 Strasbourg. CEDEX, France
EDRS PRICE	MF-$0.83 Plus Postage. HC Not Available from EDRS.
DESCRIPTORS	Adult Education; *Adult Students; *Educational Accountability; Educational Assessment; Evaluation Methods; Foreign Countries; Language Instruction; Language Proficiency; *Language Programs; *Language Tests; Listening Comprehension; Reading; *Second Language Learning; *Self Evaluation; Student Evaluation; Verbal Ability; Writing Skills
IDENTIFIERS	Finland; France; Israel; *Sweden

ABSTRACT

This study was designed to outline possible forms of guided self-assessment in adult language learning and to suggest steps that might be taken in order to further develop methods judged to be suitable for use within the European unit/credit scheme for foreign language learning by adults. The report includes a review of relevant literature and research activities, a survey of practices in various educational institutions in Scandinavia and elsewhere, concrete suggestions for self-assessment materials, and a description of field experiments with those materials. Additionally, proposals for further research and practical development work are made. Research activities in Sweden, France, Israel, and Finland are outlined. In March 1977 a questionnaire requesting information about self-assessment practices was sent to 120 Swedish educational institutions and organizations providing foreign language instrucion for adults. The most common form of self-assessment, apart from the self-assessment element inherent in most types of structured exercises with model answers, consisted of ordinary paper-and-pencil tests that the learner scores himself with the help of an overlay stencil or an answer key. A guidance test for university and college students and use of self-instructional and self-assessment materials in research projects in Sweden are discussed. Suggestions on self-assessment materials concern standardized forms, assessment in relation to individual goals, standardized forms based on threshold level, self-assessment by means of formal tests, and informal self-assessment. Self-assessment forms are included. (SW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

COUNCIL OF EUROPE

CONSEIL DE L'EUROPE

ED166962

"PERMISSION TO REPRODUCE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY
Council of Europe
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.

Council for Cultural Co-operation
COUNCIL OF EUROPE
Strasbourg
1978

APPROACHES TO SELF-ASSESSMENT IN FOREIGN LANGUAGE LEARNING

by

Mats Oskarsson

Language Teaching Research Centre
University of Göteborg
Sweden

TABLE OF CONTENTS

Page

1. INTRODUCTION 1

1.1 Introductory remarks on the role of evaluation within the unit/credit system 1

1.2 Aims of the study 3

2. RELATED RESEARCH 4

2.1 Introduction 4
2.2 Sweden 4
2.3 France 5
2.4 Israel 7
2.5 Finland 8

3. SURVEY OF PRACTICES 8

3.1 Questionnaire and interviews 8
3.2 Self-assessment by means of formal tests 8
3.3 Guidance tests for university and college students 9
3.4 Use of self-instructional and self-assessement materials in research projects 11

4. CONCRETE SUGGESTIONS 11

4.1 Introduction 11
4.2 Self-assessment by means of standarised forms 12
4.21 Global assessment 12
4.22 Concrete situations 14
4.23 Descriptive rating scales 15
4.3 Assessment in relation to individual goals 18
4.31 Individual goals and individualised instruction 18
4.32 Self-assessment in relation to specific learning objectives 19
4.4 Standardised forms based on threshold level objectives 20
4.41 Topics 20
4.42 Functions 21
4.5 Final Note on self-assessment forms 22
4.6 Self-assessment by means of formal tests 23
4.61 Assessment of general language proficiency 23
4.62 Assessment of specific language behaviours 23
4.63 Self-administration of threshold level tests 24
4.64 Formal tests based on the threshold level specifications of words and structures 24

./.

Page

4.7 Informal self-assessment 25
4.71 Informal self-assessment in a group 25
4.72 Informal self-assessment in authentic situations 26

5. EXPERIMENTS 26

5.1 Practical application 26
5.2 Experiments with self-assessment forms (types B and C) 27
5.3 Experiments with other self-assessment forms 29
5.4 An experiment with C_1 forms and a self-assessment form based on the threshold level specification of topics 31
5.5 Follow-up experiments 32

6. FURTHER RESEARCH AND DEVELOPMENT WORK 32

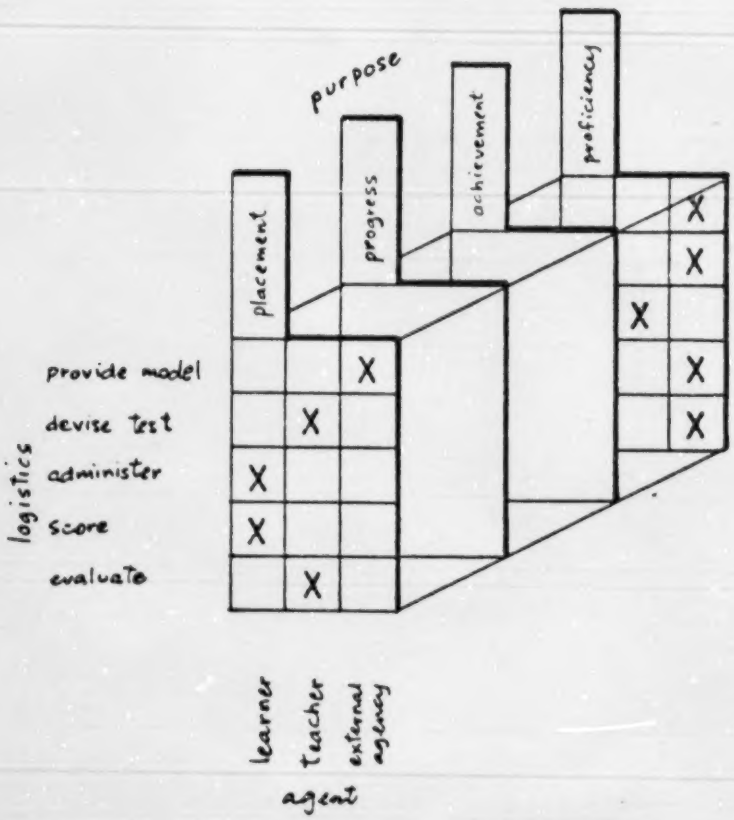
Bibliography 35

Appendices 39

PREFACE

This study forms part of a broad programme of work, embracing a wide variety of interrelated contributions, which is now rapidly developing within the Modern Languages Project of the Council of Europe. The project has grown out of the work of a group of experts set up in 1971 by the Council for Cultural Co-operation of the Council of Europe. This expert group was charged with the task of formulating proposals for the establishment of a European unit/credit scheme for foreign language learning by adults. The group has consistently based its work on two premises. First, that the provisions made for language learning should be seen as constituting a coherent system , and that any particular aspect or component (eg needs analysis, specification of objectives, course design, teaching methodology, evaluation) should be viewed in this context. Secondly, that at every stage the intentions and resources of the learner should be the controlling factors for reaching proper decisions as to what he should learn, and how he should learn it.

If this is true for the choice of goals and methods, it is also true for evaluation: the point of testing is to give information to the student, and also to others for their mutual benefit, about his standing in relation to his appropriate learning. The range of possible evaluations is wide, from an apparently unassailable, objectively validated grade (pass or fail) with international currency to a momentary expression (smile or frown) on the face of a teacher. The roles of parties involved vary according to situations and requirements, as illustrated in the following diagram.



One area which needs to be explored and cultivated in the jungle of possible forms of evaluation is how the individual student may be brought to an efficient realisation of his own achievement, or lack of it, in relation to the goals he set himself in the first place. This is one kind of evaluation which is particularly important for the adult learner, who may well be highly motivated, but needs this motivation reinforced by practical applications and feedback at frequent intervals, preferably timed according to his own demands. This requires that he should have available to him not only means of assessing his own performance, but also means of learning how to make the necessary assessments. Dr. Oskarsson's study is an important first step, an exploration of several possible forms of self-assessment for use in adult language learning. The author's intention is, as he says, " to suggest lines along which further research and development work seems to be most likely to yield fruitful results, rather than to present detailed and final material for ready use". It is in this spirit of exploration and discovery that the expert group invites readers of this study to report back on any experience they may gain with the forms of self-assessment suggested in it as well as with other forms and procedures which may help students to use their learning experiences, to improve their performance and to apply the language they have learnt to European communication and understanding.

J L M Trim
Project Adviser

Correspondence concerning this study, including the reproduction or translation of all or part of it, should be addressed to the Director of Education and of Cultural and Scientific Affairs of the Council of Europe, F-67006 Strasbourg-CEDEX.

1. INTRODUCTION

1.1 Introductory remarks on the role of evaluation within a European unit/credit system for modern language learning by adults

It has been recognised that evaluation within the unit/credit system will require many and varied testing techniques. An early symposium on, "Linguistic content, means of evaluation and their interaction in the teaching and learning of modern languages in adult education" (Rüschlikon, May 1971) reached the conclusion that "As the mastery of a foreign language, even at a fairly low level, is too complex a matter for evaluation to be based only on objective techniques, *both objective and subjective assessment must be employed*" /our italics/ (Council of Europe document, 1971, p.4). In a paper presented on the same occasion Kingsubry (1971) remarked that "If assessment within units is to be truly effective, then most certainly *a multiplicity of assessment techniques will have to be employed* /our italics/. If a system of continuous assessment is applied, one must recognise that continuous assessment is applied, one must recognise that continuous assessment does not mean simply a series of formal tests given at irregular intervals" (p. 6). At the symposium on "A European unit/credit system for modern language learning by adults" at Ludwigshafen in 1977 the participants jointly recommended that "effective evaluation tools, ranging from self-assessment instruments to standardised tests for defined objectives" (Council of Europe document, 1977, p. 4) be developed.

Accordingly, we will assume that a wide range of evaluation procedures and test types will be needed in order to achieve the "multi-faceted learning system adapted to their (ie the learners') individual needs and objectives" (Trim, 1973, p. 1) which is the project's aim.

The importance of involving the learner in the question of evaluation has also been repeatedly emphasized in Council of Europe documents dealing with the unit/credit project:

> "It was felt strongly that the student should be given a greater responsibility for determining the pace of his own learning and should be involved wherever possible in the process of evaluation" (Conclusions, Rüschlikon Symposium; see Council of Europe document, 1971, p. 4).
>
> "With regard to the European unit/credit system for language learning, it is essential to find an appropriate instrument to evaluate the communicative ability of the learner and help him develop a self-critical awareness of his acquisition of knowledge, so that he can find his place in the continuous learning process within the framework of a unit/credit system" (Conclusions, St. Wolfgang Symposium; see Council of Europe document, 1973, p. 9).
>
> "In an adult education system, student control of the training process should be extended to evaluation procedures, and exercised in different ways according to the functions - whether educational, scientific or social - which evaluation is expected to perform" (Leon, 1976, p. 1).

./.

"The education of the learner to make reliable and valid autonomous judgements on the effectivenss of his communication is a necessary part of the learning process" (Council of Europe document, 1976, p. 7).

"A European language policy of this kind should be designed so as to serve the basic ideals of the Council for Cultural Co-operation such as:

...

to make the process of language learning more democratic by providing the conceptual tools for the planning, construction and conduct of courses closely geared to the needs, motivations and characteristics of the learner, and enabling him so far as possible to steer and control his own progress" (Trim, 1977, p. 1).

Engaging the learner actively in the evaluation of learning effects will probably lead to greater interest in techniques for continuous assessment, as opposed to terminal or "end-of-unit" assessment, which is emphasised in traditional testing. This does not mean that assessment by the learner himself is entirely different in purpose from formal outside testing. The general trend towards self-management in adult education is sometimes taken to include evaluation aspects as well, so that even decisions concerning adequate initial standards for a particular field of study may be made by the learner himself on the basis of self-administered tests (cf for example Section 3.3).

In prcinciple, self-assessment can thus be both terminal (or summative in the terminology of Bloom, Hastings, and Madaus, 1971) and continuous, although it is very seldom practised for the former purpose. By tradition, terminal evaluation is what we have formal tests and central testing institutions for, as it were. Assessment by the learner himself has typically been regarded as synonymous with "formative exercises" belonging to the student's practice book. It is only quite recently, as a result of the increased demands for needs-oriented, learner-centred education, that a new attitude towards langugage testing has begun to emerge. There is a growing feeling that tests ought to be integrated into the learning porcess, chiefly in order to provide feedback on progress to the learner, and that they ought to be controlled, where feasible, by the learners to a greater extent than is now the case. Rivers (1973) expresses this standpoint in the following words:

"Instead of better comparative /ie norm-referenced/ tests, we should seek for a better system of establishing criteria consistent with the student's personal aims and purposes, a system in which the student himself would be involved, so that his progress would be clear, at any point, and continuous. The test as an extra activity then becomes largely superfluous, unless students request it for the pleasure of demonstrating their achievement. When students enjoy tests as a challenge and an opportunity for displaying what they know, we shall have reached the optimal form and timing of the test" (pp 31-32).

./.

This was probably written with college and university students in mind, but it does apply in the field of adult language learning as well. Many, if not most, adults who participate in courses do so because they want to satisfy their desire to learn something, not because they need, or attach great importance to, a certain formal attestation of their performance. Most of those who pursue foreign language studies purely for their own personal satisfaction usually do not have general education in their own language beyond compulsory school requirements and ordinarily they have no plans to raise their formal qualification levels by embarking on a comprehensive further education "venture". Moreover, as has been pointed out by Gorosch (1970), for instance, "Adults who have reached a certain position in life, are rather sensitive to formal evaluation in language proficiency, particularly that concerning oral skill, knowing that the latter is rather difficult to aquire with age" (p. 19).

The most important function of tests and evaluation within the unit/credit system, and indeed in any learner-centred language learning system, is consequently to provide constant diagnostically oriented feedback to the learner as he tackles the various components, or units, of the course work. In this perspective, terminal evaluation for control or accreditation purposes is only of secondary importance.

As the quotations above show, there is fairly common agreement among experts on the need for a wide variety of continuous assessment techniques, but the problems involved have not yet been solved. Discussing this question, Clark (1972) remarks that "Although most proponents of individual instruction seem to appreciate the role and necessity for intensive formative testing in connection with individualised programmes, little substantial progress has been made yet toward resolving a number of practical measurement problems inherent in such an undertaking" (p. 227). Others have warned against the dangers that may lurk in the continuous assessment approach if applied unjudiciously. Gorosch (1971), for instance, expresses concern for the possible mental stress that might result from continuous assessment, because "it might well cause the same nervous pressure, but on more occasions than the traditional system of final examinations" (p. 11). It might also be difficult to combine self-instruction and continuous assessment, he feels.

In summarising, it may be said that a wide variety of assessment techniques will obviously be needed in order to accomplish the flexible learner-centred system which is foreseen in the unit/credit project. It also seems clear that self-assessment of the results of learning is going to be an increasingly important feature of the system and that it is necessary to make provision for the development of evaluation techniques which can be put in the hands of the learners. The most important function of such self-assessment techniques ought to be to give the individual continuous feedback on what he has learnt. Ideally self-assessment procedures should also enable the learner to assess his total achievement at the end of a course or course unit.

In this study, the emphasis is not on continuous self-assessment, since the necessary framework for experimentation in this field, ie a concrete threshold level course, has not yet been developed. In the opinion of the group of experts on evaluation, however, priority should be given to work on continuous assessment practices as soon as such a course becomes available.

1.2 Aims of the study

The aim of the present study is to outline possible forms of guided self-assessment in adult language learning and to suggest steps that might be taken in order to further develop methods judged to be suitable for use within

the unit/credit project. The report includes a review of relevant literature and research activities, a survey of practices in various educational institutions in Scandinavia and elsewhere, concrete suggestions for self-assessment materials, and a description of field experiments with those materials. Finally proposals for further research and practical development work are made.

2. RELATED RESEARCH

2.1 Introduction

A survey of literature relevant to self-evaluation of foreign language proficiency has been undertaken. Most of what has been done in the field involves self-administration of regular tests. Attempts at using more unconventional techniques have also been made. Some of them seem to hold out a great deal of hope for further successful work on self-assessment.

The following account is natrually far from exhaustive. It merely exemplifies some activities that seem to have a direct bearing on our own efforts.

2.2 Sweden

In connection with an investigation of the validity of recently developed English tests for adults (Balke-Aurell, 1977), self-assessment results and formal test results were compared with independent, impressionistic assessments made by native English and American language teachers and, in some cases, by native instructors working in close professional contact with the subjects. The self-assessment instrument consisted of a questionnaire setting out a variety of sociolinguistic situations. The subjects were requested to rate their linguistic ability to cope with each situation on an 11-point scale. They were also asked to assess their general proficiency in each of the four language skills, also on an 11-point scale.

The teachers and the instructors rated the subjects by means of a questionnaire consisting of 60 statements concerning various aspects of language behaviour. The statements represented different levels of performance. The "assessors" (teachers and instructors) were also asked to indicate, on 11-point scales, the subjects' general proficiency level in each of the four language skills. These four items were identical with the ones the subjects themselves answered.

In all, 266 subjects employed in two different sectors of industry (commerce and technical engineering) took part in the study. All of them answered the self-assessment questionnaire and afterwards took the proficiency test battery, which comprised listening comprehension, reading comprehension, writing proficiency, and oral proficiency (only 40 subjects took the latter part; cf below). 85 subjects were rated by a teacher or a native English or American instructor. 40 subjects were rated by two native instructors independently of each other, and, in addition, also took the oral production part of the test battery.

./.

Several analyses were made, including intercorrelation between self-assessment scores, ratings by the assessors and formal test scores.

It was found that there was a very high intercorrelation between the instructors' independent assessments of the same subjects (around .90 in each of the four language skills), and that the subjects' assessment of their own language proficiency correlated around .60 with the assessment of the instructors. This means that the instructors rated the subjects in very much the same way and that there was fairly good agreement between the instructors' opinions and the subjects' own opinions. The author concluded that both validation instruments appeared to work well as criterion measures.

Formal test results correlated no higher than self-assessment scores with the instructors' judgements (coefficients ranging from .40 to .60). Self-assessment scores and formal test results correlated around .50. It is difficult to say which of the two criteria, instructors' judgements or formal test scores, is the most accurate. We only know that the tests were quite comprehensive and very well tried out, and that the instructors based their judgements on extensive personal contacts with the subjects.

Since there are several elusive variables involved here, one must of course be careful in an interpretation of the results. However, the general character of the experiment seems to warrant the conservative conclusion that global self-assessment techniques of the impressionistic or introspective kind need not be very much inferior to formal testing procedures in respect of general information yielded.

2.3 France

CRAPEL. One of the most important characteristics of the unit/credit project is its endeavour to produce a learning system which puts the learner and his needs in the foreground. The idea is of course not new and has been practised in many places. At the Centre de Recherches et d'Applications Pédagogiques en Langues (CRAPEL) of the University of Nancy, for instance, the notion of student-centred learning is one among several areas which have received particular attention (cf for instance Cembalo and Holec, 1973).

We will not go into the details of the CRAPEL learning philosophy, but a few guiding principles in the autonomous scheme, relevant to our aims, are worth mentioning.

A distinction is made between teaching and learning. The focus of interest is on the learning process, not on the activities of the teacher. Indeed, in CRAPEL terminology there are no teachers, only "helpers" whose function it is to facilitate learning. The helpers' most important task is to make the learners aware of the different approaches to language learning available in order to enable them to make an informed personal choice of goals and methods. This means that the learner who so wishes can study the language in a traditional kind of learning situation with a conventional teacher-student relationship and text-centred methods of teaching. More often, though, the learner opts for a more or less autonomous working model in which part of the "lessons" are characterised by truly self-directed learning on an individual basis or in a small unsupervised group. Helping the learner to "learn how to learn" by himself is considered a very important objective. Or, as a member of the CRAPEL staff, Philip Riley (1976) puts it, "a teaching strategy which respects the special conditions of the

adult learner will be one which enables him to teach himself and to be himself" (p. 18). Generally the role of the helper is played down. The evaluation by the learner himself of the learning process is considered important.

The ultimate goal is to make the learner "truly autonomous". This means that he will have acquired "the ability to define his objectives, his working conditions and the content and method of working, as well as to evaluate and adjust his progress. To meet these conditions, ... , an increasing number of activities are introduced during the non-systematic /ie semi-autonomous/ stage which are directed at the acquisition of pedagogic autonomy. These include group sessions devoted to the topics which have been indicated, as well as practical demonstrations of different kinds of exercises, exchange of documents and information and so on" (Riley, 1976, p. 20).

The importance of authentic language materials and language situation is emphasised. The learner is encouraged to bring along materials of particular interest to him and, at higher levels of proficiency, to establish contacts with native speakers living in Nancy. Each year CRAPEL draws up a list of the names and addresses of Englishmen and Americans who are willing to meet with the learners for "conversation hours". It is the learner's responsibility to take the initiative, and to make the arrangements for a meeting. The use of language for oral communication is stressed. Native teachers/helpers and authentic language materials (recordings of television and radio programmes, newspapers, letters, etc) help to simulate real "communication" situations which motivate the learner to acquire communicative competence. The emphasis on the functional aspects of language learning is of course very much in line with the intentions behind the unit/credit scheme.

Assessment in the autonomy scheme is completely in the hands of the learner in the same way as "the notion of teacher competence has been revised and the final responsibility for learning left to the learners themselves" (Henner Stanchina, 1976, p. 74). Helpers and course organisers do not use any kinds of formal tests in order to evaluate learning results, since that would be contradictory to the expressed educational philosophy. Instead evaluation consists of "the learners' more or less subjective views of the savoir-faire they have acquired and the perception of their attainment of self-established goals" (Abe, Henner Stanchina, and Smith, 1975). No formal procedures for self-assessment have been developed; the main tool is thus the learner's introspection, his own observation of his performance in authentic situations, eg when meeting a native speaker, when reading a professional journal, when listening to a radio broadcast in the foreign language, or when talking to his matched peer (see below). The key question the learner is advised to ask himself is: Can I perform to my own satisfaction? If the answer is no, which of course is the normal case, then the learner will have to decide whether he needs some sort of evaluation of his standards, and, if so, what form such evaluation should properly take.

Objective views on language proficiency may be solicited from the helper or from native speakers who have had conversation hours with the learners. It is emphasised that such "independent estimates" cannot be considered final in any way, since communicative competence (which is what

./.

most of the learners are striving to attain) is determined by many factors which the learner himself is in the best position to judge. Among these are of course purely psychological characteristics such as the ability to adjust to new situations and to interpret and effectively use para-linguistic signals for communication (signs, gestures, etc).

The initiators of the CRAPEL autonomy scheme naturally realise that it may be difficult for learners to evaluate their own progress and performance objectively. "Depending on their ability to adopt self-critical attitudes, some will under-rate, others will over-estimate themselves, and in either case the uncertainty as to the moment when learnering can be said to have taken place remains" (Henner Stanchina, 1976, p. 38). The best indication of the learner's linguistic competence is of course real verbal interaction in the meetings with native speakers that are offered by CRAPEL. Before the learner has reached the stage where he or she is ready to meet a native speaker, communicative ability can usually be tested in conversation with a matched peer. This peer-matching system, whereby learners who are at approximately the same proficiency level are introduced to each other, also provides opportunities for the learner to measure his or her performance against that of someone else. Since autonomy to some extent is "synonymous with a certain dose of solitude", it is essential to give learners this "opportunity to meet with each other, and to confront real situations (authentic documents, native speakers) which will provide the feedback necessary for more objective, more constructive evaluation" (Henner Stanchina, 1976, p. 38).

CRAPEL has extended a generous offer of co-operation with our project concerning further research on self-assessment (personal communication with H Holec and C Henner Stanchina; see further Chapter 6).

UNESCO. In a study at the Ecole Normale Supérieure de Saint-Cloud, France, commissioned by UNESCO, a method has been developed for determining the degree of autonomy of a given programme of study (Candelier et. al. 1975 and 1976). The means used is a large set of self-scoring evaluation items which, when completed, allows the "self-learner" to compare his programme with other learners' programmes. The evaluation sheets make it possible to rate self-teaching methods on, among other things, the extent to which they are appropriately adapted to autonomous study, or, in project terms, to autodidaxy (total autonomy) or semi-autonomy.

It should be noted that this study was directed at the evaluation of learning programmes, not learning itself; however, certain items concerned the evaluation of learning (essentially whether tests are determined in advance and whether part of them can be taken without the mediation of a teacher).

2.4 Israel

In a study at the Hebrew University of Jerusalem, Berkoff (1975) administered objective, semi-objective, and subjective tests of reading comprehension, vocabulary and composition writing to 50 advanced students of English. They were asked to evaluate their performance on the tests and the self-estimates were then compared with the actual scores and the

./.

teachers' evaluation. (The complete date of the study was not available at the time of writing.)

2.5 Finland

Tampere University, Finland, reports experiments with tests of spoken English in group situations (Folland, 1976). The testees are given a stimulus from a prerecorded tape which presents a topic and a range of discussion points. Following this presentation a group discussion takes place during which all testees participate. The examiner uses a special marking sheet where he assesses each testee's linguistic performance in relation to the number of utterances, how effectively language is used for communication etc.

According to the author the assessment "can be relatively objective, is technically simple and demands less of the examiner than the usual interview-type test". If this is so, similar techniques could possibly be developed for use by the learners themselves.

Tampere Language Institute reports on experiments with what is called "foreign language conversation testing" (Bacon and Ojanen, 1976). The testee is given a choice of three topics and is then allowed 15 minutes in which to prepare a 3-5 minute talk on one of the topics. There then follows a discussion between the testee and two native speakers on the topic chosen. The assessors award points from one to five (from "As native" down to "Errors affect understanding") in relation to four different factors. "The system is not difficult to operate" according to the authors and might therefore be a possible form of self-assessment in autonomous groups.

3. SURVEY OF PRACTICES

3.1 Questionnaire and interviews

In March 1977 a questionnaire requesting information about self-assessment practices was sent out to 120 Swedish schools, institutions, and organisations providing foreign language instruction for adults. Information has further been gathered through interviews with representatives of various adult education institutions. Several experienced folk high school teachers, study circle leaders, and teachers in municipal adult education have given their opinions on testing in general and have outlined various practices within their respective areas.

Various techniques are reported, such as self-administration of formal tests, evaluation in groups, evaluation by means of closed-circuit television and videotape recordings, etc. The methods used are mostly very informal and not well documented.

3.2 Self-assessment by means of formal tests

The most common form of self-assessment, apart from the self-assessment element inherent in most types of structured exercises with model answers, consists of ordinary paper-and-pencil tests which the learner scores himself with the help of an overlay stencil (transparency) or an answer key. The

./.

procedure is quite commonly used by adult education organisations which offer courses at various levels and who wish to ensure that the learners are placed at levels appropriate to their actual learning stages.

For practical reasons such tests are mostly restricted to the measuring of word knowledge, mastery of grammar, and reading comprehension, usually by means of conventional written tests. The measurement of listening comprehension and oral and written proficiency as a rule requires somewhat more elaborate testing tools and is therefore seldom included in self-evaluation procedures.

The tests are usually referred to as placement tests or, in most cases misleadingly, diagnostic tests.

In order to simplify further the evaluation of tests results some institutions and organisations use self-scoring answer sheets. Figure 1 illustrates a commercially available type consisting of two sheets with a double-sided carbon-paper in between. The set is sealed at the edges and can easily be opened after use. Keys showing the correct options are printed on the back of the top sheet (Teacher's Key) and on the bottom sheet (Student's Key). As the student marks an answer it is copied on both keys.

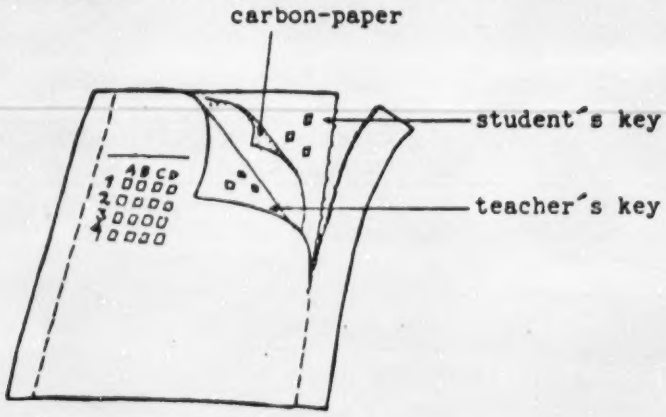


Figure 1

Self-Scoring Answer Sheet for Multiple Choice Tests

When used for formal testing purposes, the student hands in the completed set to the teacher/test administrator, who then opens it, gives the student the bottom sheet (Student's Key) and keeps the top sheet (Teacher's Key) for marking. By this procedure the student gets immediate feedback on his performance and the teacher can mark and evaluate the test at his own convenience.

3.3 Guidance test for university and college students

As of July 1, 1977, Sweden introduced new and more liberal requirements for admittance to universities and other kinds of higher education. Among those now free to enroll are persons belonging to "the 25:4 category", ie people at least 25 years of age who have worked for at least four years. The only "general requirement" is a certain, very moderate level of achievement in Swedish and in English as a foreign language. In order to make it possible for those considering a university education to find out for themselves whether their proficiency in English is sufficient, ie good

enough to enable them to study textbooks in English in their subject, the Department of Education asked the National Board of Education to produce a self-assessment test. The test was constructed and tried out within the GUME/Tests project at the University of Gothenburg during 1976-77. The test, Kan du läsa engelsk test? (Can you read English texts?) was sent out free of charge in the spring and summer of 1977 to several thousand would-be university students.

Included in the test booklet of this so-called guidance test (vägledningsprov) is a key and a section giving norms and recommendations. Those whose results fall below certain specified levels are recommended to study English prior to taking up full-time studies, especially if their subject is one in which many of the required books are in English. (There are subjects where all the texts are in Englsih.) The minimum level, which is fairly low, was decided on by the Department of Education in relationship to what students in the upper secondary school (gymnasieskolan) with grades slightly below average produce. The test had been tried out in the upper secondary school; the level is norm-referenced rather than referenced to a criterion of actual reading performance in a university setting.

The test takes a maximum of 2 1/2 hours to do. It can be ordered free of charge from the GUME/Tests project, University of Gothenburg, S-431 20 Mölndal, Sweden. All those who have ordered the test have been asked to make a note of their results and to be prepared to answer a questionnaire which will be sent out towards the end of the spring term of 1978 in order to establish the predictive validity of the test. A couple of hundred students in a number of departments at the University of Gothenburg have also taken the test under controlled conditions at the beginning of the autumn term of 1977 and will be similarly approached a year later and asked to describe how they have succeeded in reading English textbooks.

The investigation will, hopefully, result in at least partial answers to questions like the following: To what extent did the students take the advice offered in the recommendations (in other words, does the principle of self-assessment work when it comes to decisions which have far-reaching implications for the individual)? How well, or how poorly, did students with low scores do in their subsequent studies?

A group which has been given special attention in the area of self-assessment are the trainees at the colleges of education who train to teach forms 1 through 6 of the nine-year comprehensive school. They are also going to teach English themselves in forms 4-6. In the late 60's it was found that many students finished their three-year course of study without being able to get passing grades in English and a need was felt for a self-assessing prognostic test. Such a test was first produced by the GUME/Tests project at the request of the National Board of Education in 1972. New versions and new tests have been used in succeeding years by almost 10,000 students. The screening effects of these tests have proved to be good: many have decided not to start their training at all, others have postponed starting for a year. Those who have decided to start irrespective of poor test results have usually failed but have at least been warned early on in their studies.

./.

In 1977 a new test was produced, called Engelskt basfärdighetsprov (English Basic Skills test), in which the command of elementary grammar and vocabulary is tested at a level rather below that which the average pupil in the 6th year is able to produce. This test is also accompanied by recommendations for those who score very low.

A similar test was produced in the spring of 1977 for teachers who teach immigrant children their various native languages (they are referred to as hemspråkslärare, "home language teachers"). This vocabulary and reading comprehension test in Swedish as a foreign language, Pröva din svenska! (Test your Swedish!), was also produced within the GUME/Tests project at the request of the National Board of Education in order for teachers to ascertain whether they would be able to take part in teacher training conducted in Swedish (dozens of languages being represented among the teachers).

All the above tests were produced for adults planning to take up higher education of various kinds in which some knowledge (mostly on what might be referred to as an intermediate level) of a foreign language was required. As far as can be judged, the tests have turned out to function very well for these purposes.

3.4 Use of self-instructional and self-assessment materials in research projects

Self-instructional and self-assessment materials have been used in various ways and for various purposes in research projects. As an example may be mentioned experiments concerned with problems of foreign language learning in adult education (von Elek and Oskarsson, 1973). Adult learners were taught "grammar units" by means of programmed materials which included prerecorded lessons played back in class from tape recorders, sets of projector transparencies for over-head screen display, and student workbooks which contained dialogues, illustrations, and exercises. For each "teaching point" covered there was a variety of tasks and self-tests which the subjects completed and then scored and evaluated individually with the help of a key projected on a screen in front of the group. In some cases the correct responses were explained and commented on after the subjects had been given time to make corrections in their workbooks. On an average the "grammar units" comprised two lessons each and were offered according to two alternative modes of learning.

The programmes were also adapted for use with cassette recorders in individualised instruction (von Elek, 1975).

This "step-increment" method of programmed learning and continuous self-assessment proved to work very well as a formal pre-test/post-test comparison showed. The progress rate was quite substantial and student assitudes were generally favourable.

4. CONCRETE SUGGESTIONS

4.1 Introduction

Broadly speaking, self-assessment is a new field in language testing and consequently there is very little accumulated knowledge and experience to draw on for our purposes. Therefore this section of the report is necessarily of a tentative and provisional nature. It is intended to suggest lines along which further research and development work seems to be most likely to yield fruitful results, rather than to present detailed and final

material for immediate use by others. Failing solid experience on which to base our appreciation of the techniques proposed, further elaboration and empirical experimentation will certainly be needed.

Possible ways of introducing the notion of self-assessment in the regular school setting will on the whole be left out. Suffice it to say that self-assessment in its simplest form consists of informal discussions between teacher and pupil about performance and grades. The pupil can for instance be asked to indicate his relative position in the class or in the group according to his own estimate (cf quotation from Upshur, 1975, in Section 4.2), or he can be asked to describe what he feels to be his language skill in concrete operational terms.

In principle, the methods outlined on the following pages could also be used in the compulsory school system, but they have primarily been devised with adults in mind.

A few rather unconventional evaluation techniques will first be outlined. In Chapter 5 an account will be given of some small-scale experiments with some of the proposed techniques. It ought to be pointed out that an attempt has been made to find some sort of a basis for decisions about possible new directions of research in this feidl and therefore the emphasis is not on the obvious and quite common procedures which involve self-administration of formal tests (cf however Section 4.6). There are basically two reasons for this. One is that the unit/credit system aims at being innovative and at contributing new and alternative approaches in the field. The other is our concern for the learner's possibilities of assuming more responsibility for the structuring and evaluation of his own learning. We feel this can best be achieved if we do not rely on traditional testing methods only.

4.2 Self-assessment by means of standarised forms

4.21 Global assessment

One way of giving the learner a better basis for his own assessment of language skills is to provide him with various kinds of standarised forms which rely on the learner's ability to make over-all impressionistic and introspective judgements. As has been observed from time to time, even objective, well-constructed and standarised language tests seldom tell the whole truth, just as it can be said that even the most subjective estimate always contains some truth. Discussing subjective and objective aspects of language testing Upshur (1975) reports a case where it was found that the best way to measure someone's oral proficiency was simply to ask him how well he spoke the language. The author comments:

> "This is not very surprising. I think most of us could give a pretty accurate appraisal of our own abilities in other languages that we know. And certainly it is the rare student is a small audiolingual class who cannot tell which of his classmates are poorer and which are better than he is" (p. 329).

./.

We will start, therefore, by looking at some subjective assessment forms which are based on the hypothesis that most learners have a certain capacity for determining their own language ability - provided they have at their disposal a measuring standard by which they may express their intuitions.

A model of a very simple standard form is illustrated in Figure 2. It consists of a ten-point scale (or a scale with any arbitrary number of points) ranging from complete mastery of the skill in question down to no ability at all. Only those two "anchor points" are described verbally. The learner is instructed to put a cross in the box which he feels corresponds to his proficiency level. The text is given in the learner's native language (ie German for native German learners, French for native French learners, etc) in order to avoid mistakes due to insufficient knowledge of the target language.

		Put a cross in the box which corresponds to your proficiency level.
LISTENING COMPREHENSION	I understand practically everything that is said to me in English.	10
		9
		8
		1
	I do not understand spoken English at all.	0

Figure 2

Example of Self-Assessment Form (Listening Comprehension), type A

A set of four forms of this kind, one for each of the four language skills, could be used as a point of departure for discussions about the learner's command of the language. A set of guiding principles would of course have to be spelled out. The discussions could be between the learner and his teacher or between the learner and his fellow students. In both cases the procedure will result in valuable information to the learner.

./.

Naturally enough this particular kind of form can only yield very rough estimates of stages of learning and cannot be used for assessment of learning in very restricted contexts. On the other hand, it is no doubt feasible to adapt the model to suit specific learning objectives as well. Instead of assessment in relation to the global concept "Listening comprehension" one might for instance use "Understanding an English news broadcast on the radio", "Reading an English daily paper" or even "Re-telling a story recently read in the textbook".

It would be desirable to experiment with self-assessment materials of this kind in order to determine their usefulness. In the opinion of some experts very little faith can be put in measuring instruments of this type. However, only controlled experiments can give us an indication of the extent to which they can be usefully employed for evaluation purposes.

A set of four Type A forms is included in Appendix A.

4.22 Concrete Situations

To the Type A non-defined scale one can also supply descriptions of concrete situations each of which requires mastery of a particular language skill. An example of such a form (Self-Assessment Form, Type B) is shown in Figure 3. Again the text should be in the learner's native language, not in the language the learner is studying.

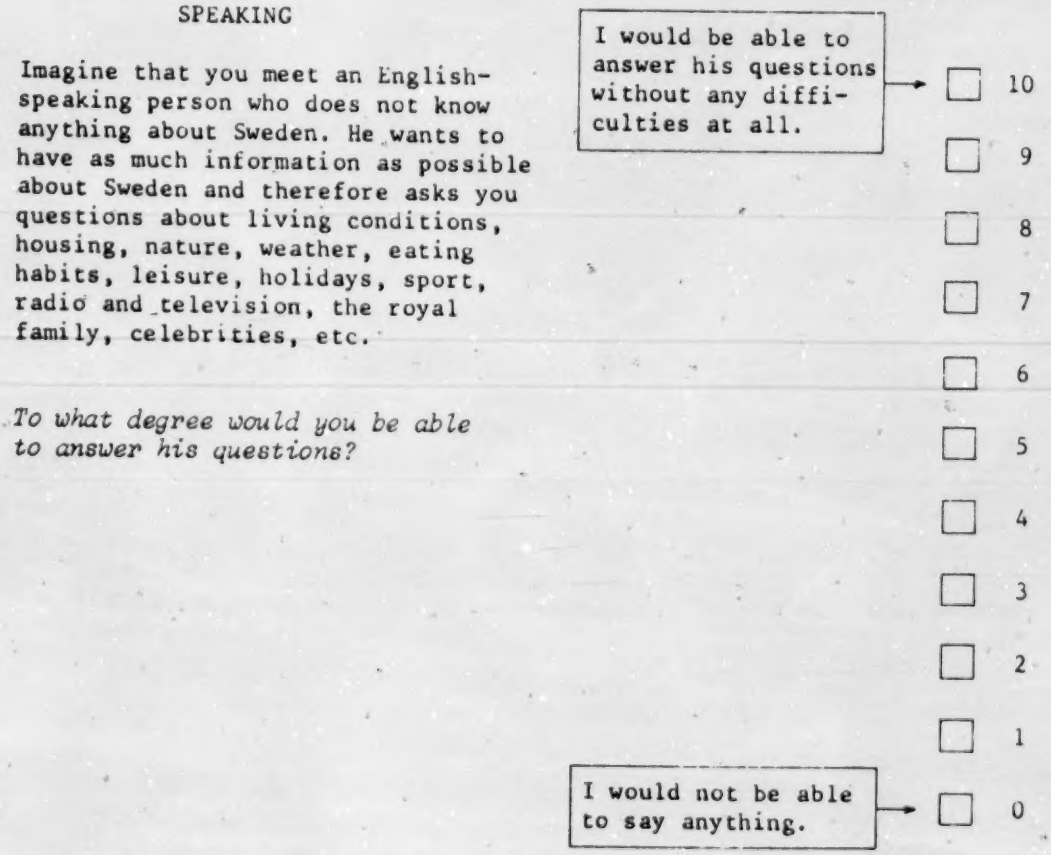

SPEAKING
Imagine that you meet an English-speaking person who does not know anything about Sweden. He wants to have as much information as possible about Sweden and therefore asks you questions about living conditions, housing, nature, weather, eating habits, leisure, holidays, sport, radio and television, the royal family, celebrities, etc.
To what degree would you be able to answer his questions?
I would be able to answer his questions without any difficulties at all.
10
9
8
7
6
5
4
3
2
1
I would not be able to say anything.
0

Figure 3

Example of Self-Assessment Form (Speaking), Type B

./.

Box 10 in these scales would correspond to native proficiency and Box 0, of course, to no proficiency at all. The Threshold Level as it is now conceived could probably be placed at either Box 2 or Box 3, and the Waystage Level at 1 or 2.

The situations chosen should require as little specialised factual knowledge as possible, so as to ensure that the response does not unduly reflect the learner's "knowledge of the world" rather than actual linguistic competence.

As is all too well known in the profession, one thing that mars language teaching today is the fact that it is so notoriously difficult to produce valid and practicable oral proficiency tests (for a discussion of the problems involved, see Clark, 1972). The continuing shift in emphasis towards communicative objectives has highlighted the problem. In view of this, and in view of the fact that the unit/credit system is conceived of as essentially a system for the development of aural/oral skills, it might be worthwhile making particular efforts to develop this administratively simple approach to measuring oral communicative competence. It would seem to be possible to link the description of concrete situations quite closely to threshold level objectives, for instance. A student who has studied a Threshold Level course could evaluate himself in relation to the behavioural specification of Topics (van Ek, 1975, Chapter 8) in this manner, using, for instance, item 5.3, which says that the learner should be able to "characterise countries, places and regions, say something about their size, where they are situated and in what surroundings, mention possibilities for sight-seeing; seek similar information from others, what to visit and how". Empirical data on where students who have reached Threshold Level place themselves on the scale, on an average, would be the norm against which the individual student checks his own estimate.

Appendix B contains a specimen form for each of the four skills. They have been used for experimental purposes in a few groups of adult students (see chapter 5).

4.23 Descriptive Rating Scales

In most cases it would be more fruitful, and easier for the learner, to use forms in which all points on the scale are defined in concrete terms. A non-language-specific and comprehensive form of this kind, including all four language skills, has been worked out and used in pilot experiments (Self-Assessment Form, Type C); cf. Chapter 5. It is illustrated in Figure 4 (one skill only). This type of form, too, ought to be written in the learner's native language.

./.

Form for

SELF-ASSESSMENT OF LANGUAGE PROFICIENCY

Instructions: Put a cross in the boxes corresponding to your level in each of the four skills (LISTENING, SPEAKING, READING, WRITING). Only one cross for each skill!

SPEAKING	
☐ I speak the language as well as a well-educated native.	5
☐ I speak the language fluently and for the most part correctly. I have a large vocabulary, so I seldom have to hesitate or search for words. On the other hand I am not completely fluent in situations in which I have had no practice with the language.	4
☐ I can make myself understood in most everyday situations, but my language is not without mistakes and sometimes I cannot find the words for what I want to say. It is difficult for me to express myself in situations in which I have had no opportunity to practise the language. I can give a short summary of general information that I have received in my native language.	3
☐ I can make myself understood in simple everyday situations, for example asking and giving simple directions, asking and telling the time, asking and talking about simpler aspects of work and interests. My vocabulary is rather limited, so it is only by a great deal of effort that I can use the language in new and unexpected situations. I am sometimes misunderstood because of faulty pronunciation and using words wrongly.	2
☐ I can just about express very simple things concerning my own situation and my nearest surroundings, for example asking and answering very simple questions about the time, food, housing, and directions. I only have a command of very simple words and phrases.	1
☐ I do not speak the language at all.	0

Figure 4

Example of Self-Assessment Form (Speaking), Type C
(first version)

./.

With the help of such aids it ought to be possible for the learner to form a rough, or maybe even a fairly accurate opinion of his level of learning. The complete set of forms, comprising all four language skills, of course makes it possible to define an approximate learning profile as well.

The forms were used in a few experiments (cf. Section 5.3) and were then modified to include non-defined half-steps between the defined levels. The complete revised forms are given in Appendix C.

The specification of levels in any scale of this kind is problematic. The more levels there are, the more difficult it is to define each step. For our particular purposes (ie self-evaluation by unsophisticated learners) it was not judged possible to distinguish clearly, in easily understood terms, between more than six levels (including the zero and full proficiency levels). An experienced evaluator, on the other hand, may be able to use a more elaborate scale comprising a larger number of defined levels.

The content of the descriptions is of course crucial. Ideally different learners should understand them in the same way, but this cannot be taken for granted. From a psychological point of view it would further be preferable to exclude negative definitions such as "sometimes I cannot find the words for what I want to say" in favour of consistent positive wording, but then the explicitness would probably suffer. It is considerably easier to establish a distinct hierarchy of levels if one is allowed to use both positive and negative statements.

It might be possible to link this kind of descriptive scale to formal tests, such as the Threshold Level Test, by establishing rough equivalences between scale levels and test scores. In order to do this one would need to administer the self-assessment scale together with the test to a sizeable group of students representing different stages of learning. By such a procedure it would further be possible to relate scale levels to other defined levels of performance. Some work of this kind has been carried out in Sweden; rough comparisons between scale levels and stages in the Swedish school system can be made.

With slight modification these self-assessment scales could be used for the rating of others or as a general framework in language teaching contexts. Instead of using the first person singular in the defining statements ("I can ..." etc), impersonal constructions would be used.

It has to be emphasised, however, that the primary aim of the present scales is to provide learners with a global self-assessment tool. For other purposes there are of course other scales available as well (cf for instance, Trim 1977, in which John Trim also discusses the general usefulness of generalised performance ratings).

One final but very important point ought to be made. It concerns the need for training. In order to be successful in the use of subjective self-assessment materials of the kind discussed here most learners would need some practice in the use of the techniques involved. Otherwise there

./.

is a risk that personality factors may influence, or even invalidate, the results obtained (see further Section 4.5). Proper training in self-observation and in making judgments ought to be made part of the learning process.

4.3 Assessment in Relation to Individual Goals

4.31 Individual Goals and Individualised Instruction

One of the main purposes of the planned unit/credit scheme is to develop a learner-centred and flexible instructional system which can cater to individual needs:

> "A unit/credit system of this kind appears to offer a promising framework for guidance combined with freedom and flexibility in the at present badly understructured field of adult education.
>
> The aim of the European unit/credit system for modern language learning by adults is therefore to establish a framework for adult language learning, based upon the language needs of the learner and the linguistic operations required of him in order to function effectively as a member of the language community for the purposes, and in the situations, revealed by those needs" (Trim, 1973, p. 18).

If these expectations are met in the planned unit/credit system, we will have at our disposal, for the first time really, a means enabling us to provide each adult learner with a truly individualised instructional programme which at every point is clearly related to personal learning objectives. The advantages of such an individualised system over conventional "lockstep" teaching are attested to by most educators today. A broad range of arguments to this effect can be found in the ACTFL Review of Foreign Language Education, Vol. 2 (Lange, 1974), System (1977), Altman and Politzer (1971), and Valette and Disick (1972). (The latter book, by the way, contains self-tests based on its contents.)

The shift of emphasis from the teaching process itself to individual differences in modes and strategies of learning, to differences in learning goals, and to behavioural objectives, is accompanied by a redirection of interest from norm-referenced tests, ie tests where the learner's performance is matched against that of other learners, to criterion-referenced tests, where the learner's performance is evaluated in relation to some predetermined criteria, ideally concrete language behaviours (cf for instance Steiner, 1974). The latter type of test has the advantage that "matching against the criterion becomes a challenge and a guide to further effort" (Rivers, 1973, p. 28) instead of "a hurdle to be surmounted". Rivers goes on to say that in using criterion-referenced tests "the student knows exactly what knowledge he must demonstrate and either demonstrates it and moves on or cannot demonstrate it and goes back to see how he can improve his performance". The explicitness of performance objectives of the kind embodied in the unit/credit model likewise makes for greater clarity and "gives all those involved in the teaching/learning process, including the learner himself, a clear view of just what is expected of them" (van Ek, 1975, p. 22).

./.

Criterion-referenced tests and performance objectives thus facilitate the adaptation of instruction to individual goals, and the learner can assume more responsibility for his own training, so that "when he feels ready to match the criterion, he tests. Note that he tests: it is not we who test him. If necessary, he later retests. When satisfied that he has matched the criterion, he moves on" (Rivers, 1973, p. 28). This spirit squares well with the overall philosophy of the unit/credit system ("a learner-centred, needs and motivation-based language learning system",; Trim, 1976, p. 78) as well as with the notion of self-assessment.

In an article dealing with the need for objective evaluative criteria in language teaching Valdman (1971) points out that criterion-referenced tests are an integral part of programmed self-instructional materials. If the unit/credit system is going to be realised in a fashion resembling programmed instruction, which conceptually and Practically is inextricably associated with individualised learning, we will of course be able to draw on the experience gained with stimulus-response drills and other rigorously structured exercises on discrete learning points. As has been pointed out, such tasks can in fact be regarded as being made up of test items each of which offers instant feedback to the learner on his performance level.

Others have expressed concern for "the dehumanising effect of evaluative procedures in individualised programmes" (Jorstad, 1974, referring to Giorgiades, 1973).

The general trends in adult language learning, ie greater emphasis on performance objectives, individual learning, and criterion-referenced, learner-centred evaluation, are represented in the unit/credit approach and ought to be reflected in the kinds of assessment techniques applied as well.

4.32 Self-Assessment in Relation to Specific Learning Objectives

For a majority of adult learners in the Council of Europe member countries the evaluation of general, non-specialised language proficiency is probably more important than the evaluation of highly specialised language abilities. Most learners who seek to evaluate their performance will want to know, in the first place, what stage they have reached generally in the language, or else, what stage they have reached in relation to a certain language course they have studied. This does not mean, of course, that the evaluation of achievement within a well-defined area, or in a very specific part of a course of study, is uninteresting. On the contrary, for a good many learners such evaluation is very useful and necessary as an integral part of their learning (cf for instance Valdman, 1971). Knowledge of one's gradual progression is always conducive to the attainment of established goals, or, as the case may be, to the realistic adjustment of one's goals. Evaluation of this kind, ie continuous assessment which helps the learner to form a clear conception of how he is progressing towards his personal objectives, should therefore be considered as a very essential feature of any major educational programme.

./.

The kinds of methods proposed so far can be best used for global evaluation purposes. In most cases this means some sort of terminal assessment of learning. It should be pointed out, however, that the methods are not necessarily unsuitable for intermittent use, for example during a course of study which extends over a year or two. Most learners who work seriously and for some duration of time will notice gains even on these global self-assessment instruments.

4.4 Standardised Forms Based on T-Level Objectives

Forms of self-assessment, and particularly forms of continuous self-assessment, within the unit/credit system cannot be specified in great detail until it has been made possible to outline more clearly how the system is going to be implemented in actual learning programmes. That is, very detailed proposals can only be made when the nature and contents of concepts like unit and module have been given more precise definitions.

> "The construction of a system for evaluation should be closely integrated with the building up of a unit/credit system and be based on the content of the units" (Conclusions, Rüschlikon Symposium; see Council of Europe document, 1971, p. 5)

The most concrete specifications of unit/credit objectives so far are to be found in "The Threshold Level" (van Ek, 1975) and "Un niveau-seuil" (Coste et. al., 1976), and therefore they will be the main points of reference in the discussions which follow. The English document, being the one most widely known to date, will be the primary souce of reference.

At every stage of the development of self-management techniques it is of fundamental importance to keep the target-group learner in mind. Materials and instructions must be easy to understand and easy to handly, otherwise they may not function in the hands of those who have little academic training. This constraint prevents us from sampling freely from the T-level documents when we want to translate.

4.41 Topics

One sort of objectives specification that seems to lend itself quite readily to adaptation to easily comprehensible self-evaluation procedures is the behavioural specification of topics in "The Threshold Level" (van Ek, 1975, pp. 22-28). This list specifies what the learner who has reached Threshold Level will be able to do with regard to each of the topics included in the definition of the Threshold Level (pp. 14-16). As an example we may take Topic 1 (Personal identification) under which heading we find, among several others, the entry name. The behavioural specification with regard to this topic states that the learners should for instance be able to spell out their names if required. This is a very concrete performance objective which of course can be tested in various ways. Furthermore, most learners would have no difficulty in telling whether they could measure up to this requirement or not. The simplest way to find out would therefore seem to be to ask the straightforward question: Can you spell your name in English? or, in declarative form:

I can spell my name in English ☐ Yes ☐ No

./.

The average Threshold Level learner would check the Yes box, the average learner below Threshold Level would check the No box. A range of statements (a check-list) of this kind therefore ought to be a helpful guiding tool for the learner who is working towards the Threshold Level objectives.

Working on this assumption I devised a self-assessment form comprising 30 statements representing different topics in the Threshold Level specification (see Appendix D). The form was then used in a very modest try-out with a group of adult learners (cf Section 5.4).

Again it should be pointed out that this kind of evaluation instrument is primarily intended for autonomous learners and others who have an interest in finding out, for themselves, their levels in relation to Threshold Level objectives but who do not wish to, or who cannot, take a formal, centraliy administered test. It is suggested that a guideline for evaluation be supplied at the end of the form (cf Appendix D):

> "If your total number of Yes crosses is 25 or above, and if your judgment of your own language ability is fairly accurate, you are likely to have reached Threshold Level in English".

This number, 25 positive responses out of 30 possible, is a subjective estimate which of course would have to be empirically verified.

Note also that the test should be given in the learner's native language.

Instead of forcing the learner to choose between the absolute Yes/No dichotomy, it might be more rewarding to allow for some latitude of response, eg by complementing with an 'uncertain' box or by using a scale with several steps (for instance 'yes', definitely 'yes, probably', 'uncertain', 'no probably not', 'no, definitely not'). Compare for example the attitude scales in the Appendix of Richterich and Chancerel (1977).

Once a fully specified Threshold Level course, or any unit/credit course, becomes available, it will no doubt be possible to work out similar aids for continuous self-assessment of progress. Such forms would then be modelled to fit in at successive points during the course.

4.42 Functions

Another possible starting-point for discussions about self-assessment within the unit/credit framework is the specification of language functions and exponents of those functions that are found in "The Threshold Level" and "Un niveau-seuil". As will be recalled a distinction is drawn in 'The Threshold Level between six main categories of language functions, viz. 'imparting and seeking factual information', 'expressing and finding out intellectual attitudes', 'expressing and finding out emotional attitudes', 'expressing and finding out moral attitudes', 'getting things done (suasion)', and 'socialising' (van Ek, 1975, pp. 19-21). Under each of these headings there is a further specification of functions such as 'expressing agreement and disagreement', 'accepting an offer or invitation', 'inquiring whether something is considered possible or impossible', etc. The document also contains a list

./.

of the actual language forms (grammatical and/or lexical) which are considered to be the most useful exponents for each function (ibid. pp. 34-44).

It appears that these kinds of objectives specification could be easily comprehended by the learners if combined and then adapted for self-assessment purposes in the same way as was done with the topics specification above. To illustrate, we may take the function 'inquiring about agreement or disagreement' (ibid. p. 19). The specification of language forms (exponents) lists the following items as representative of the function (ibid. p. 36):

Do you agree?
Do you think so, too?
Don't you agree?
Don't you think so?
(short questions)

In self-evaluation form this particular item might read as follows:

I can ask a person if he or she is of the same opinion as I am. (To be formulated in the learner's native language.) ☐ Yes ☐ No

Examples: Do you agree?
Don't you think so?

An attempt has been made to draw up a provisional self-assessment instrument representative of the inventory of language functions found in "The Threshold Level" (see Appendix E). It has not yet been possible to test it in a field situation, but research activities are being planned (cf Chapter 6).

In "Un niveau-seuil" (Coste et. al. 1976) the list of functions and their most common exponents in French ("Actes de paroles"/"Actes d'ordres", pp.830224) occupies a central position. The section would seem to be well suited to the same kind of adaptation for self-assessment techniques as was proposed for the English counterpart.

4.5 Final Note on Self-Assessment Forms

To conclude this section it ought to be said that we do not ignore or underestimate the technical unreliability that effects any assessment procedure based on impressionistic estimation. Nor do we overlook the fact that unpredictable personality variables such as varying degrees of self-confidence, perceptiveness, and general ability to form valid judgments may be important factors. Naturally none of these traits co-varies with linguistic ability and therefore we must acknowledge that in certain extreme cases the learner will get nothing out of using purely subjective assessment methods. On the other hand it is probably fair to say that it is only a small minority of learners whose abilities to perceive and discern deviates so markedly from "the norm" that their judgement of foreign language standards is of no value at all. Secondly, evaluation by means of subjective questionnaires is only seen as one variety of self-assessment

./.

procedure. Others include assessment in autonomous groups or between peers, self-observation in authentic language situations, self-tests through exposure to mass-media (radio and television broadcasts, newspapers, tapes, cassettes, records, etc), the use of self-administered conventional tests with keys (see Section 4.6). Impressionistic self-assessment by means of questionnaires etc ought to be seen as a complement to such methods.

Our idea, therefore, is not to substitute self-assessment procedures for objective formal tests, but rather to aim at an optimal combination of the two approaches. Both have their merits - self-assessment for being learner-centred and potentially capable of taking the aural/oral aspects into full account, and formal tests for being more reliable (at least in the technical sense of the word) and better known. It therefore seems that our aim should be to try to strike a happy balance between the two.

4.6 Self-Assessment by Means of Formal Tests

4.61 Assessment of General Language Proficiency

As was pointed out earlier, the most common self-assessment aid, apart from the opportunities for tacit self-observation that pattern drill types of exercises and programmed self-instructional materials offer, consists of conventional written tasks accompanied by a list of correct answers which the learner goes through after having completed all the items. Especially at the beginning stages of learning such tasks need to be of the multiple-choice type, so that no uncertainty as to the correctness of any particular answer arises. Open-ended questions and free production tasks are more difficult to supply with simple answer models. Often it is quite impossible to anticipate all possible and correct responses; usually some responses are more correct, or more wrong, than others (which motivates a system of varying values). However, fill-in-the-blank and other variants of short completion tasks can be used to the extent that the test constructor can predetermine alternative, acceptable responses.

4.62 Assessment of Specific Language Behaviours

Discrete-item multiple-choice tests can confidently be said to measure general language proficiency quite well, provided they are prepared and evaluated in a professional manner. When it comes to the diagnostic measurement of detailed behavioural objectives they are not quite as adequate, however. As has been explained by Clark (1972), part of the reason for this lies in "the difiiculty of identifying within a given item a specific linguistic feature whose knowledge by the student is both necessary and sufficient for a correct response" (p. 239). Another reason is the fact that in a multiple-choice of general language proficiency you are interested in the total score, regardless of what combination of items it has been obtained by, whereas in a test of discrete language behaviour you want to know whether the candidate actually masters each of the specific competencies which the individual items purport to measure. If the test is of the multiple-choice format, the chance factor may preclude any conclusions about component abilities even if the total score is relatively high. Valid tests of performance objectives should therefore represent, as faithfully as possible, the very language behaviour that the test intends to measure, ie in most cases they should require the candidate to speak and write his responses.

./.

4.63 Self-administration of threshold level Tests

Self-administration of parallel forms of the threshold level test would seem to be an obvious area for the development of self-assessment Practices within the framework of the unit/credit system. To the extent that so-called objective, or rather objectively scoreable, test types are used, one would only need to provide the learner with simple answer sheets of the sorts discussed earlier (Chapter 3). Pending the completion of The Specimen test, not much work has been carried out in this area. We may perhaps mention, however, that a draft version of the test contained several fixed-response parts, which could readily be evaluated by the learner himself.

Learner administration and self-scoring of productive (subjective) tests, for instance an oral test involving recording of responses on magnetic tape, is of course a much more problem-laden undertaking. If the learner were to be provided with a spoken model (a printed model would of course be inappropriate), he would have to compare and make distinctions between aural stimuli, which is not an easy matter for anyone at threshold level. Assistance by a fellow learner, or a group of fellow learners, might make such evaluation easier.

Likewise, application of the principle of learner-centred evaluation to tests of free productive writing creates some thorny problems, It has been proposed that writing a letter requesting information about accommodation in England be included in the English threshold level test. Even such a well-defined task could no doubt be solved in innumerable ways, which means that a handy standard model would be very difficult to produce. The feasibility of the undertaking depends of course to a great extent upon what degree of formal perfection should be expected of the learner. If the criterion is "having got the message across", no matter in what form, then the evaluation part might not be too difficult to master for a fellow learner, for instance. If, on the other hand, a certain degree of stylistic sophistication is considered important, then a fellow learner is not in a very good position to judge.

A test of oral communication ability (with emphasis on communication), eg through a structured interview (which has been suggested) or some such procedure, could theoretically be handled by a fellow learner, or by a group of fellow learners, possibly in some manner similar to those described in Section 2.5. Very great care would of course have to be taken to ensure that the necessary instructions were specific enough. It would also be helpful to compare, under controlled conditions, results obtained by learner-evaluation and results obtained by regular outside administration of the same test. The outcome could help us determine the usefulness of this kind of peer-evaluation. It is therefore proposed that the follow-up activities of this project include some research on this point.

4.64 Formal Tests Based on the Threshold Level Specifications of Words and Structures

In Section 4.4 above we tried to outline some unconventional approaches to evaluation based on the topics, functions and exponent categories set out in "The Threshold Level". Both the English and the French threshold level documents (van Ek, 1975, and Coste et. al., 1976, respectively) also contain

specifications of the words and grammatical structures that are most likely to be employed in T-level courses. Although the advisability of such a measure may be rightly questioned, it is of course quite possible to devise formal self-assessment tests on these words and structures. With proper instructions, and in combination with other techniques, such tests might serve a useful purpose.

4.7 Informal Self-Assessment

The self-assessment methods that we have dealt with so far are more or less formal in nature, ie they all imply some carefully predetermined course of action or some requisite printed aid. In addition to this approach to assessment, many informal assessment methods suggest themselves. Some of these are probably more expedient and effective than the most sophisticated formal testing technique, others are only of very limited significance. Since it would be beyond the scope of the present study to try to establish a comprehensive typology of such informal assessment strategies, we will limit ourselves to commenting briefly on a few that seem to have obvious practical relevance for our purposes.

4.71 Informal Self-Assessment in a Group

The learner's assessment of his own language ability can be greatly facilitated if he is a member of a group with which he can interact. The group can assist in various ways, eg by informally discussing each member's proficiency level or progress rate (this requires a certain atmosphere of mutual trust in the group), or by providing a framework for role-play and similar communicative exercises. The group does not need to be big; even two people can be called a group for the purposes we are discussing here.

Evaluative group activities of the kinds which are our present concern should take as their starting-point the objectives stated in relevant Council of Europe documents (eg van Ek, 1975; Coste et. al. 1976). Since these objectives are generally quite explicit, it should be possible to carry out a great deal of communicative "test exercises" in accordance with unit/credit aims. In principle this would mean that the performance objectives were acted out as communication tasks. Necessary prompts could be supplied in various ways, eg by means of the native language or pictures. The participants in the communication acts would have to decide, either subjectively or by some other principle, whether the intended verbal functions had been achieved.

Autonomous interaction practice in simulated situations can range in degree of complexity from very short exchanges of greetings ("socialising"; van Ek, 1975, p. 44) to a group discussion on a topic of current interest ("expressing and finding out intellectual and emotional attitudes"; ibid. pp. 36-42). As for evaluation, most learners would no doubt be able to tell whether their own standards permitted them to interact adequately in each situation. If not, a group discussion or consultation with a fellow learner, would be helpful in most cases. Such discussions could even be turned into a semi-formal evaluation procedure if each member of the group assesses all the others, resulting in a consensus opinion for each (rather

./.

like a skating competition). For obvious reasons, this method cannot be applied without due preparation in the group. Most adults are very sensitive to openly pronounced value judgments about themselves, and therefore it is imperative that any testing activity, be it formal or informal, is expressly approved of by each individual member of the group.

Role-playing and other communication tasks should be practised continuously almost from the beginning in any orally oriented language course. As we mentioned earlier, the exact techniques to be used depend to some extent on what types of learning programmes the unit/credit system is going to embody.

4.72 Informal Self-Assessment in Authentic Situations

Authentic language situations provide the most valid opportunities for self-evaluation. It is in the actual use of the language, in real-life settings, that one may ultimately test one's communicative ability. Therefore the learner should always try to assess his ability in terms of his use of the language in natural situations, not in terms of test scores or in terms of performance in an artificial learning situation. Proper questions to ask oneself are: How well do I understand this letter, this newspaper headline, this information leaflet, this telephone message, this television interview? How well can I make myself understood when talking to somebody in the foreign language? Can I express my intentions, my feelings, my thoughts? It is less useful to ask: How many of the words in my textbook do I know? How well do I master this structural pattern? How foreign is my accent? etc.

It is hoped that further work in the field will include attempts at developing methods for self-assessment in authentic language situations.

5 EXPERIMENTS

5.1 Practical Application

Some of the ideas proposed in this project are being tested in adult education institutions in Sweden (municipal adult education centres and study circles). The notion of self-assessment along the lines presented here has been introduced for the first time in a centrally produced teachers' manual to be used at municipal adult education centres all over the country. Teacher reaction will be collected.

Field experiments have also been conducted. Briefly, the aim was to see whether there is any correspondence between teachers' judgment of their learners' proficiency and the learners' own judgment based on the self-assessment forms illustrated above. The learners were asked to make independent assessments of their proficiency in each of the four language skills. The teachers, who did not see the completed forms, were requested to provide the project with preliminary grades (on a 5-point scale) given at the time of the learners' self-assessment (about three weeks before the end of the term). On the basis of self-assessment forms only, the experimenter was then to predict the grades the teachers had given their students. The aim was to see to what extent outside evaluation and self-assessment correlated with each other. It was hoped that the results would provide a basis for the evaluation of self-assessment as a principle.

./.

5.2 Experiments with Self-Assessment Forms (Types B and C)

Three groups of adult learners at "approximately twice threshold level" took part in an experiment with Self-Assessment forms (Types B and C; see Sections 4.22-4.2 and Appendix). Each form covered all four language skills. As described above the Type B forms set out concrete situations which require mastery of particular language skills. The learner indicates his estimated ability level by putting a cross on a scale ranging from no ability at all (0) up to full ability (10). The levels between these two extreme points are not defined. The Type C forms consisted of described 6-point scales for each language skill. All points were defined in general everyday terms. here, too, the subjects were asked to check the levels which they felt corresponded to their real command of the language.

The subjects were further asked to state which grades they themselves thought they deserved. The grades, which are given on a 1-2 scale, are to reflect the learners' overall language proficiency, not any particular skills. The subjects were also asked to predict the grades their teachers were likely to decide on. Finally they were to indicate whether they had already discussed their proficiency levels and grades with the teacher.

The aim of the experiment was to see to what extent the students' self-assessment by means of standardised forms corresponded to the teachers' opinions as manifested in the course grades they intended to award the students. To achieve this I was to predict, so the basis of the questionnaires only and without having any other indication of the subjects' proficiency in the language, what grades the teachers were likely to give. The course grades were thus used as the criterion of all-round language proficiency. The teachers were not informed about the self-assessment results, and the students, with a few exceptions (see below), did not know what grades they were going to get.

Ten students participated in the first group (Group A). After I had received the forms I first calculated self-assessment "scores" (total of Forms B and C). On the basis of these total scores I predicted each student's course grade, which, as explained above, could be 1, 2, 3, 4 or 5. My judgment was also based on the knowledge that the lowest point on the scale (grade 1) is not often used.

My prediction of the teacher's grades, which I did not see until after my first estimates, was correct in 3 cases out of 10. After I had seen the students' own estimation of their "deserved grades" and their prediction of what grades the teacher was going to give them, my adjusted prediction was correct in 6 cases out of 10. The students' prediction of the grades given by their teacher were correct in 8 cases out of 10, but then two of the students reported that they had conferred with the teacher about what grades they were likely to get.

In the second group (Group B) 19 students filled in the questionnaire. One student did not answer all the items and was therefore excluded. Here my first prediction, based on the self-assessment forms (B and C) only, corresponded to the teacher's grades in 8 cases out of 18. My second estimate,

./.

based also on knowledge of the students' estimation of their "deserved" grades, and their prediction of the teacher's grades, was correct in 12 cases out of 18. Three students had discussed their grades with the teacher prior to filling in the self-assessment forms. The students' predictions of what grades the teacher was going to award was correct in 10 cases out of 18.

The third group given the B and C forms plus questions about grades consisted of 8 students (Group C). The evaluation procedure was the same as in the two previous groups. In my first round of predictions I was correct in 6 cases out of 8. After I had looked at the students' "self-assigned" grades and their expectations concerning course grades to be assigned by their teachers, my new predictions coincided with the teacher's grades in all 8 cases. Two students reported that they had talked with the teacher about their grade levels.

It needs perhaps emphasising that my predictions of grades was purely a matter of mechanical assignment on the basis of formal self-assessment scores. Certain "cutting scores" were established and applied in an uncomplicated fashion; the procedure required no expert knowledge of testing or any special language skills once the forms had been devised and the cutting scores had been decided on. The calculations could very easily have been made by the learners themselves. All that would have been needed was a set of very simple instructions to follow. It would probably take the average learner no more than a quarter of an hour to evaluate himself/herself along the steps involved in the present experiment.

The results of the three experiments are summed up in Table 1:

Table 1

Proportion of Correct Predictions of Course Grades in Self-Assessment Experiments

	On the Basis of Self-Assessment Forms (Types B and C)	On the Basis of Self-Assessment Forms and Self-Evaluated Grades	Students' Predictions of Teachers' Course Grades
Group A	$\frac{3}{10}$ (30%)	$\frac{6}{10}$ (60%)	$\frac{8}{10}$ (80%)
Group B	$\frac{8}{18}$ (44%)	$\frac{12}{18}$ (67%)	$\frac{10}{18}$ (56%)
Group C	$\frac{6}{8}$ (75%)	$\frac{8}{8}$ (100%)	$\frac{6}{7}$ (86%)
Total	$\frac{17}{36}$ (47%)	$\frac{26}{36}$ (72%)	$\frac{24}{35}$ (69%)

On the basis of the information provided by the students' self-assessment forms I could thus assign, in a strictly formalised manner, the students to the "correct" grade category in 17 cases out of 36. As indicated above I worked on the assumption that the grades the teachers award the learners are true reflections of all-round language skills. It may of course be the case that the four language skills do not carry the same weight when it comes to the evaluation of success in the courses in question.

Knowledge of the students' self-evaluated grades and their prediction of the teacher's evaluation increased by "correct" score markedly. As can be seen I was able to predict the exact grade for 26 students out of 36 (ie in nearly three cases out of four). Out of the remaining 10 I missed 9 students by one grade and only 1 by two grades. As was said before I had no knowledge of the subjects' test records or their background. The experiment was in fact conducted per capsulam, so I never even met the students.

In summary it may be said that the results of the experiment with self-assessment forms of Types B and C in three groups of adult learners seem to show that impressionistic self-assessment can yield quite informative results. The learners in our experiment had a good grasp of their actual language skills and they were able to communicate their impressions quite well by means of the evaluation instruments provided. Concrete descriptions of language behaviour, suitably devised, would therefore seem to be useful self-assessment aids in the unit/credit system.

5.3 Experiments with Other Self-Assessment Forms

A few more groups were involved in experiments, this time with a revised version of the Type C form and some other measures. In the experiments described it was found that the students occasionally put a cross midway between two levels in the descriptive rating scales (Form C). In order to make provision for the need of a more finely graded instrument, half-steps were added to the original scales. The total number of levels was thus increased to 11. As before only the levels of whole numbers were defined. For reasons stated earlier (Section 4.23), I did not try to specify the intervening levels in concrete verbal terms.

./.

The revised descriptive scale is illustrated in Figure 5 (one skill only):

Form for

SELF-ASSESSMENT OF LANGUAGE PROFICIENCY

Instructions: Put a cross in the boxes corresponding to your level in each of the four language skills (LISTENING, SPEAKING, READING, WRITING). Only one cross for each skill!

WRITING		
☐	I write the language as well as a well-educated native.	5
☐		4.5
☐	I write the language rather easily and for the most part correctly. I only make occasional grammatical mistakes and spelling mistakes. When writing ...	4
☐		3.5
☐		1.5
☐	I can formulate very simple messages connected with my own life. I only have a command of very simple words and phrases.	1
☐		0.5
☐	I cannot write the language at all.	0

Figure 5

Illustration of the Revised Descriptive Scales for Self-Assessment (Form C_1)

The complete descriptive scales for self-assessment are given in Appendix C.

./.

Languages other than English were also involved. A small group studying German (Group D) filled in the four revised descriptive scales (Form G). As explained earlier, the use of the scales is not confined to any particular language.

Only 5 learners were present at the meeting when the forms were administered. On the basis of the self-assessment scores I anticipated the teacher's final grades correctly in 2 cases. After knowledge of self-estimated grade levels and students' predictions of the teacher's grades my predictions were correct in 3 cases out of 5. The remaining 2 cases were missed by 1 grade level each. The students who ranked highest on the self-assessment scales were also going to get the highest grades, according to the teacher.

In another small group (Group E: N=4) studying French at approximately Niveau-seuil the learners' self-estimates by means of the C_1 forms corresponded to the teacher's impressions fairly well in three out of four cases. The fourth learner clearly underrated her real proficiency in French by 1 to 1.5 scale levels in all four language skills. It should be observed, however, that this group used only a single global assessment instrument (ie the C_1 forms).

It is not uncommon to find that adult learners, especially middle-aged and older, underestimate their achievement in this way. An informal comparison between Groups A-C, consisting mainly of young adults, and Groups D-G, consisting of adults with a higher average and tends to give a slight impression of just this kind. Therefore one must of course always try to make "only" learners in particular, but also younger adults, aware of this potential source of error.

In one group of learners of English (Group F: N = 10) self-assessment by C_1 forms was compared with results on a formal paper-and-pencil test covering vocabulary, grammar, reading comprehension, and listening comprehension. The correlation between the two measures was only moderate, but then it must of course be remembered that the self-assessment forms (C_1) are more global in nature than the formal test, which, for instance, did not take the oral production skills into account.

The formalised assignment of the learners to grade categories was successful, again in terms of correspondence with the teacher's grades, in 4 cases out of 10. In general, those who ranked themselves highest were also the ones who received the highest grades. Adding information about results on the formal test increased the success rate to 6 out of 10.

5.4 An Experiment with C_1 Forms and a Self-Assessment Form based on the Threshold Level Specification of Topics

The experiments described so far have all been directed towards global self-assessment of language ability. This was done in order to find out whether self-assessment can be considered reliable enough for further development in the context of testing in the unit/credit system. As we have seen the learners' appreciation of their own standards may be more realistic and accurate than is sometimes thought. It therefore seems worthwhile to proceed to more restricted learning situations, such as are envisaged within the unit/credit system, for instance, in order to find out whether the principle of self-management in testing works there as well.

To achieve this end a group of adult learners studying English just above threshold level (Group G) were given the revised C forms plus the self-assessment form based on the threshold level behavioural specification of topics (Self-assessment Form Type D; see Section 4.41). For practical reasons the two kinds of forms could not be administered on the same occasion.

When the C_1 forms were filled in, 11 people were present; four of these absent at the second meeting, when the Topics form was administered. As in the previous experiments the teacher's grades were also collected.

Assigned grade levels on the basis of C_1 self-assessment forms corresponded to the teacher's grades in 6 cases out of 11. For the 7 learners who filled in the Topics questionnaire (Form D) a renewed prediction was made; this time the added information tended to confuse the picture rather than contribute to a better approximation of the final grade distribution. Only 2 of the 7 learners were placed in correct grade categories, 2 learners who were first placed in their final grade categories were misplaced because of their responses to the Topics form.

The result of this very limited trial should not be taken as an indication that the Type D forms, specifying topic behaviours, is unsuitable as a basis for self-evaluation. It would be ill-advised to discard it without further empirical evidence. The number of learners in the group was too small, for one thing. Furthermore, the general proficiency level was above rather than at or below T-level, which had the effect that most of the subjects checked nearly all of the Yes boxes (the mean number of Yes responses was 26 out of 30 possible). This so-called restriction of range of coures severely limited the instrument's chances to distinguish between individuals, and therefore it ought to be tried out in groups within the T-level range.

5.5 Follow-up Experiments

From the experiments conducted with groups A-F it seems clear that standardised forms of Types B-C, involving self-evaluation in relation to concrete linguistic situations and a descriptive rating scale, respectively, can be useful aids in autonomous evaluation. However, the experiments reported can only be regarded as very modest pilot activities needing much more elaboration. It is therefore proposed that further experiments be undertaken in order to obtain more conclusive evidence (cf Chapter 6). Other self-assessment techniques, including group assessment and assessment in pairs, should be incorporated in such research.

6 FURTHER RESEARCH AND DEVELOPMENT WORK

The present project has met with considerable interest among learners, teachers, administrators and research workers. Several institutions have expressed their interest in participating in future research and development work and provision has already been made for such co-operation. The CRAPEL at the Université de Nancy, which is particularly interested in developing methods for autonomous language learning (cf Section 2.3), has

./.

offered to provide groups in different countries. The offer has been gratefully accepted and plans are being made for the administration of experimental material through the CRAPEL. The Pädagogische Arbeitsstelle of the Deutscher Volkshochschul-Verband, Frankfurt/Main, where much significant work in adult language testing has been undertaken, is willing to share its experience with the self-assessment project. Professor Raasch at the Universität des Saarlandes, Saarbrücken, is planning to start experiments with forms presented in this report. In return self-assessment materials prepared by Professor Raasch will be administered to Swedish groups studying French. Professor Satzinger of Vienna has proposed joint try-out activities in Austria. Interest in participation has also been expressed by the Eurocentre. We are very grateful for all these offers of co-operation.

Swedish research projects and language teaching institutions and organisations have also been approached in order to attract attention to the unit/credit project and, in particular, to the field of self-assessment. Further investigations along the lines presented in this report are now in progress as part of a newly started research project at the National Swedish Board of Education. The project, called ARB-Vux, is concerned with forms of study and learning strategies in adult education. Work on self-assessment is being carried out by a sub-group representing different educational sectors (the folk high school, study circles, labour market training, and municipal adult education). Some of the ideas in the report have further been incorporated in a teachers' manual to be tried out at various adult education centres. Contacts have also been established with two study associations which are starting experiments on combinations of formal testing/self-evaluation procedures for placement purposes.

Plans are further being made for the introduction of such self-assessment techniques as have been proposed here in the context of labour market training for migrant workers (or immigrants in Swedish terminology). A certain minimum level of proficiency in Swedish is required for the various training programmes. Each trainee's level is established on the basis of ordinary paper-and-pencil tests or teacher judgments. Limited experiments with combinations of teacher evaluation, formal tests and self-assessment procedures are now under way.

It is important that provision be made for the further investigation of methods for self-assessment within the unit/credit system. It is therefore suggested that a plan for further action be drawn up, so that experiments under way and offers of co-operation can be profitably developed. Both formal and informal self-assessment strategies need to be explored in more detail than has been possible here.

./.

BIBLIOGRAPHY

Abe, D, Henner Stanchina, C, and Smith, P "New Approaches to Autonomy: Two Experiments in Self-Directed Learning". Mélanges Pédagogiques 1975, CRAPEL, Université de Nancy II.

Altman, H B, and Politzer, R L (Eds). Individualising Foreign Language Instruction. Newbury House: Rowley, Mass, 1971.

Bacon, W, and Ojanen, S L. "Foreign Language Conversation Testing at Tampere Language Institute". Kielikeskusuutisia (University of Jyväskylä, Finland), 4/1976, pp 55-59.

Balke-Aurell, G. Validering av språktest (Validation of Language Tests). Personaladministrativa Rådet, Stockholm, 1977 (Mimeo).

Berkoff, N. Discrepancies between Student and Teacher Estimates of Language Proficiency in Different Types of Tests of English as a Foreign Language. Paper read at the Fourth International Congress of Applied Linguistics, Stuttgart, 1975.

Bloom, B, Hastings, T and Madaus G. Handbook on Formative and Summative Evaluation of Student Learning. New York: McGraw-Hill, 1971.

Candelier, M, Chaix, P, Coste, D, N'Neil, C and Regourd, A. Report and Proposals for Establishing a Regional Inventory of Methods and Materials for Self-Learning of Modern Languages in Europe. Unesco, Paris, 1975 and 1976.

Cembalo, M and Holec, H. "Les Langues aux Adultes: Pour une Pedagogie de l'Autonomie". Mélanges Pédagogiques 1973, CRAPEL, Université de Nancy II.

Clark, J. "Measurement Implications of Recent Trends in Foreign Language Teaching". In Lange and James (eds), Foreign Language Education: A Reappraisal. Skokie, III: National Textbook Company, 1972.

Coste, D, Courtillon, J, Ferenczi, V, Martins-Baltar, M, Papo, E and Roulet, E. Un niveau-seuil. Council of Europe, Strasbourg, 1976.

Council of Europe documents:

Linguistic Content, Means of Evaluation and their Interaction in the Teaching and Learning of Modern Languages in Adult Education: Conclusions. Doc. EES/Symposium, 53, 12, Council of Europe, Strasbourg, 1971.

A unit/credit system for modern languages in adult education: Conclusions. Doc. EES/Symposium 57, 9, Council of Europe, Strasbourg, 197

Modern Languages in Adult Education. Doc. EES/Symposium 57, 10, Council of Europe, Strasbourg 1974.

Symposium on "A European unit/credit system for modern language learning by adults". Conclusions and recommendations. Doc. CCC/EES (77) 23, Council of Europe, Strasbourg, 1977.

./.

Engelskt basfärdighetsprov (English Basic Skills Test). Projektet GUME/prov, University of Göteborg, Mölndal, Sweden, 1977.

Folland, D. "Group Conversation as a Means of Teaching and Testing Spoken Language". Kielikeskusuutisia (University of Jyväskylä, Finland). 4/1976, pp 45-51.

Georgiades, C. "Individualised Evaluation for Individualised Learning". Thurst for Education Leadership 2 (1973), pp 39-40.

Gorosch, M. Methods and Techniques to Evaluate Linguistic Knowledge and Skills in Adult Education and their Role in Promoting Modern Language Teaching to Adults. Doc. CCC/EES (70) 124, Council of Europe, Strasbourg, 1970.

Gorosch, M. Some Current Problems Concerning Linguistic Contents and Evaluation in Modern Language Teaching/Learning for Adults. Doc. EES/Symposium 53, 2, Council of Europe, Strasbourg, 1971.

Henner Stanchina, C. "Two Years of Autonomy: Practise and Outlook". Mélanges Pédagogiques 1975, pp 73-92. CRAPEL Université de Nancy II.

Henner Stanchina, C. Autonomy: A Viable Strategy for Adult Language Learners. CRAPEL, Université de Nancy II. Master's Thesis, 1976, 97 pp.

Jorstad, H. "Testing as Communications". In Jarvis (ed). The Challenge of Communication. Skokie, III: National Textbook Company, 1974.

Kingsbury, R H. Additional Notes on the Proposed UCS in Modern Language Teaching and Learning for Adults. Doc. EES/Symposium 53, 4, Council of Europe, Strasbourg, 1971.

Lange, D L. Individualisation of Instruction (The ACTFL Review of Foreign Language Education 2). Skokie, III: National Textbook Company, 1974.

Leon, A. Note on How to Train Educators to Evaluate. Doc. CCC/EES (76) 5, Council of Europe, Strasbourg, 1976.

Richterich, R and Chancerel, J L. L'identification des besoins des adultes apprenant une langue étrangere. Council of Europe, Strasbourg, 1977.

Riley, R. "Current Trends in Language Teaching at the CRAPEL". Kielikeskusuutisia (University of Jyväskylä, Finland), 4/1976, pp 9-25.

Rivers, W. "Testing and Student Learning". In M C O'Brien (ed), Testing in Second Language Teaching: New Dimensions. ATESOL: Dublin, 1973.

Steiner, F. "Behavioural Objectives and Evaluation". In Lange, 1974, pp 34-78.

./.

System: A Journal for Educational Technology and Language Learning Systems. Vol 5, No 2, 1977, University of Linköping, Sweden.

Trim, J. A Units/Credits System for Modern Languages in Adult Education: Consolidated Report. Doc. EES/Symposium 57, 3, Council of Europe, Strasbourg, 1973.

Trim, J. Research and Development Programme for a European Unit/Credit System for Modern Language Learning by Adults. Doc. CCC/EES (73) 26, Council of Europe, Strasbourg, 1973.

Trim, J. "Languages for Adult Learners". Language Teaching Abstracts, Vol. IX, 1976, pp 73-92.

Trim, J. Report on some possible lines of development of an overall structure for a European unit/credit scheme for foreign language learning by adults. Doc. CCC/EES (77) 19-E, Council of Europe, Strasbourg, 1977.

Upshur, J. "Objective Evaluation of Oral Proficiency". English Teaching Forum, Vol. XIII, Nos 3-4, 1975, pp 328-333.

Valdman, A. "Criteria for the Measurement of Success in an Individualised Foreign Language Program". In Altman and Politzer (1971), pp 66-81.

Valette, R, and Disick, R. Modern Language Performance Objectives and Individualisation Brace Jovanovich, 1972.

van Ek, J. The Threshold Level in a European Unit/Credit System for Modern Language Learning by Adults. Council for Cultural Co-operation, Council of Europe, 1975.

van Ek, J. "The Threshold Level" in Educational and Culture: Review of the Council for Cultural Co-operation of the Council of Europe, Strasbourg, No. 28, 1975.

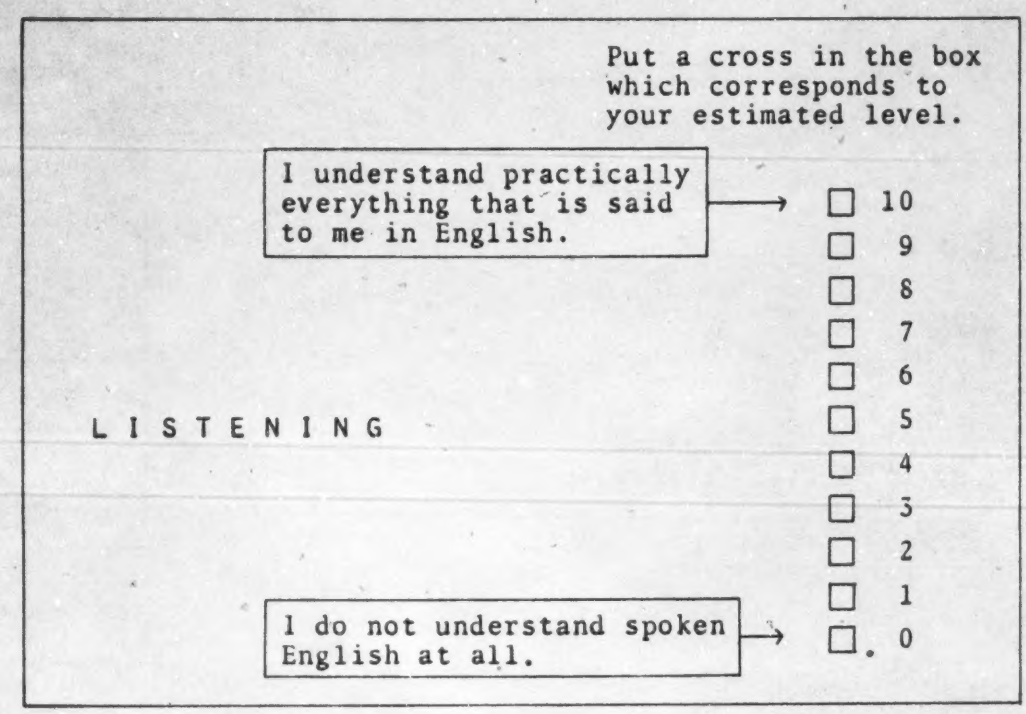
APPENDIX A

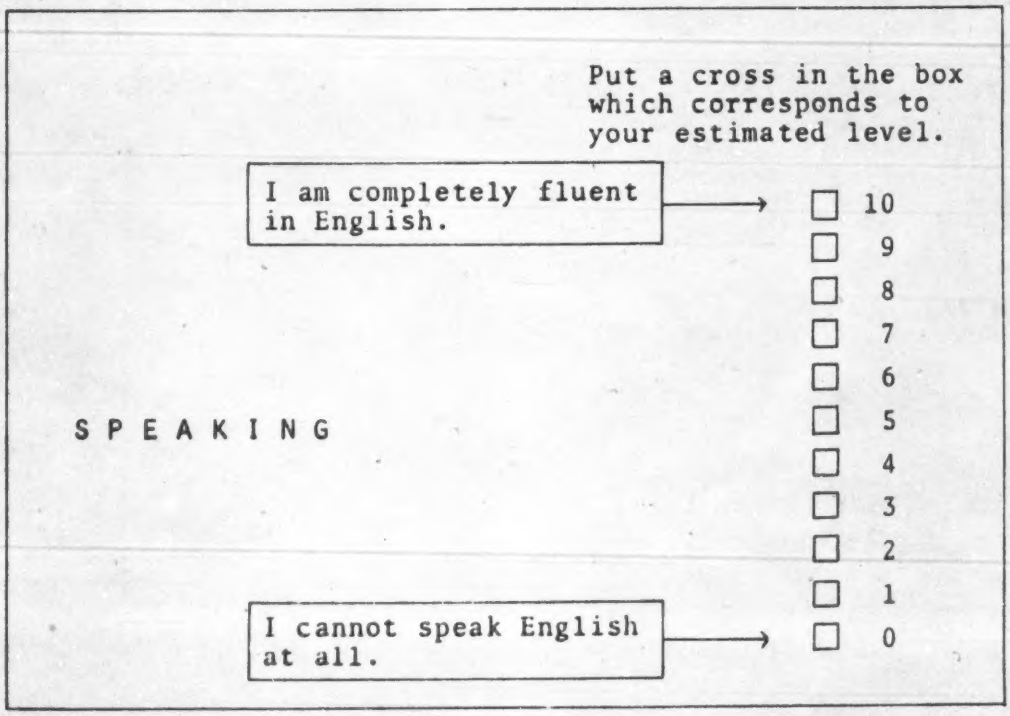
Self-Assessment Form (Type A)

SELF-ASSESSMENT FORM (Type A)

(To be translated into the learner's native language)

Instructions: Put a cross in the boxes corresponding to your estimated level in each of the four skills (LISTENING, SPEAKING, READING, WRITING).





Put a cross in the box which corresponds to your estimated level.

READING

I understand English text without any difficulty. →

- ☐ 10
- ☐ 9
- ☐ 8
- ☐ 7
- ☐ 6
- ☐ 5
- ☐ 4
- ☐ 3
- ☐ 2
- ☐ 1
- ☐ 0 ← I cannot understand English text at all.

Put a cross in the box which corresponds to your estimated level.

WRITING

I write English without any difficulty at all. →

- ☐ 10
- ☐ 9
- ☐ 8
- ☐ 7
- ☐ 6
- ☐ 5
- ☐ 4
- ☐ 3
- ☐ 2
- ☐ 1
- ☐ 0 ← I cannot write English at all.

APPENDIX B

Self-Assessment Form (Type B)

SELF-ASSESSMENT FORM (Type B)

(To be translated into the learner´s native language)

Instruction: Put a cross in the boxes corresponding to your estimated proficiency levels. N.B. Only ONE cross for each skill.

LISTENING

Imagine that you are sitting watching an episode of an English TV series such as Father Brown, The Brothers, Family at War, etc. Suddenly the Swedish sub-titles disappear.

How much would you understand of what was said if the programme continued without any Swedish text?

Everything → ☐ 10
☐ 9
☐ 8
☐ 7
☐ 6
☐ 5
☐ 4
☐ 3
☐ 2
☐ 1
Nothing → ☐ 0

SPEAKING

Imagine that you meet an English-speaking person who does not know anything about Sweden. He wants to have as much information as possible about Sweden and therefore asks you questions about living conditions, housing, nature, weather, eating habits, leisure, holidays, sport, radio and television, the royal family, celebrities, etc.

To what degree would you be able to answer his questions?

I would be able to answer his questions without any difficulties at all. → ☐ 10
☐ 9
☐ 8
☐ 7
☐ 6
☐ 5
☐ 4
☐ 3
☐ 2
☐ 1
I would not be able to say anything. → ☐ 0

READING

Imagine that someone with no knowledge of English comes to you with an English daily paper (e.g. The Times, The Guardian) and asks you to give as detailed an account as possible of what is on the front page. You have plenty of time, but must not use a dictionary.

How well would you manage the task?

I would be able to give a detailed account of everything on the page. → ☐ 10

☐ 9

☐ 8

☐ 7

☐ 6

☐ 5

☐ 4

☐ 3

☐ 2

☐ 1

I would not be able to say anything. → ☐ 0

WRITING

Imagine that you are given the task of writing a detailed account about yourself in English. It would include your life from when you were small up to the present day (childhood memories, games, school, friends, youth, present occupation, interests, personal qualities, relatives and friends, memorable experiences, future plans etc.) You have plenty of time, but must not use a dictionary.

How well would you manage the task?

I would have no problems in writing a detailed account which was linguistically correct. → ☐ 10

☐ 9

☐ 8

☐ 7

☐ 6

☐ 5

☐ 4

☐ 3

☐ 2

☐ 1

I would not be able to write anything. → ☐ 0

APPENDIX C

SELF-ASSESSMENT FORM (Type C)

(Descriptive Rating Scales)

Form for

SELF-ASSESSMENT OF LANGUAGE PROFICIENCY

Instruction: Put a cross in the boxes corresponding to your level in each of the four skills (LISTENING, SPEAKING, READING, WRITING). Only one cross for each skill!

LISTENING		
☐	I understand the language as well as a well-educated native.	5
☐		4.5
☐	I understand most of what is said in the language, even when said by native speakers, but have difficulty in understanding extreme dialect and slang. It is also difficult for me to understand speech in unfavourable conditions, for example through bad loudspeakers outdoors.	4
☐		3.5
☐	I can follow and understand the essential points concerning everyday and general things when spoken normally and clearly, but do not understand native speakers if they speak very quickly or use some slang or dialect.	3
☐		2.5
☐	I can follow and understand the essential points concerning everyday and general things when spoken slowly and clearly, but in the course of conversation I often have to ask for things to be repeated or made clearer. I only understand occasional words and phrases of statements made in unfavourable conditions, for example through loudspeakers outdoors.	2
☐		1.5
☐	I understand the meaning of simple requests, statements and questions if they are spoken slowly and clearly and if I have a chance of asking for them to be repeated. I only understand common words and phrases.	1
☐		0.5
☐	I do not understand the language at all.	0

SPEAKING	
☐ I speak the language as well as a well-educated native.	5
☐	4.5
☐ I speak the language fluently and for the most part correctly. I have a large vocabulary so I seldom have to hesitate or search for words. On the other hand I am not completely fluent in situations in which I have had no practice with the language.	4
☐	3.5
☐ I can make myself understood in most everyday situations, but my language is not without mistakes and sometimes I cannot find the words for what I want to say. It is difficult for me to express myself in situations in which I have had no opportunity to practise the language. I can give a short summary of general information that I have received in my native language.	3
☐	2.5
☐ I can make myself understood in simple everyday situations, for example asking and giving simple directions, asking and telling the time, asking and talking about simpler aspects of work and interests. My vocabulary is rather limited, so it is only by a great deal of effort that I can use the language in new and unexpected situations.	2
☐	1.5
☐ I can just about express very simple things concerning my own situation and my nearest surroundings, for example asking and answering very simple questions about the time, food, housing and directions. I only have a command of very simple words and phrases.	1
☐	0.5
☐ I do not speak the language at all.	0

READING	
☐ I read and understand the language as well as a well-educated native.	5
☐	4.5
☐ I understand everything or nearly everything written in the language within non-specialised fields. There may be words I do not understand in difficult texts.	4
☐	3.5
☐ I understand most of what I read in simple texts dealing with familiar subjects such as leisure interests, current affairs and living conditions. I understand most of a normal private letter dealing with everyday things such as the family and their doings etc. I understand the main contents of a normal newspaper article about a plane crash or the opening of a new underground line, for example, but not all the details.	3
☐	2.5
☐ I understand the meaning of simple written instructions about the way, time, place and similar things, and also understand the essential things in simple texts dealing with familiar subjects such as common leisure interests, current affairs and living conditions.	2
☐	1.5
☐ I understand the main points of a simple text and simple written directions for familiar things.	1
☐	0.5
☐ I cannot read the language at all.	0

WRITING	
☐ I write the language as well as a well-educated native.	5
☐	4.5
☐ I write the language rather easily and for the most part correctly. I only make occasional grammatical mistakes and spelling mistakes. When writing about subjects in which I have had no opportunity to practise the language I need to use a dictionary.	4
☐	3.5
☐ I can formulate written messages or give a coherent account of things connected with my own life, my interests, needs and wishes, but I make mistakes in both grammar and spelling. I often cannot find the words for what I want to express. I can write down from dictation a normal prose text about a familiar subject. I may make spelling mistakes and mistakes due to a lack of words.	3
☐	2.5
☐ I can formulate simple written messages connected with my own life and my needs, but there are often grammatical and spelling mistakes and a wrong choice of words. I can write down from dictation a simple text about everyday subjects. There are often spelling mistakes and mistakes due to a lack of words.	2
☐	1.5
☐ I can formulate very simple messages connected with my own life. I only have a command of very simple words and phrases.	1
☐	0.5
☐ I cannot write the language at all.	0

APPENDIX D

SELF-ASSESSMENT FORM (Type D)

Sample Questionnaire for
SELF-ASSESSMENT AT THE T-LEVEL

(to be translated into the learner's native language)

Instruction: Imagine that you meet an English-speaking person from another country. He does not know anything about you and your country. Indicate your estimated command of the language by putting a cross in the appropriate box (Yes or No) for each statement.

---

1. I can tell him when and where I was born.	☐ Yes	☐ No
2. I can spell my name in English.	☐ Yes	☐ No
3. I can describe my home to him.	☐ Yes	☐ No
4. I can tell him what kinds of food and drink I like and don't like.	☐ Yes	☐ No
5. I can tell him about my interests (hobbies, interests in general etc.)	☐ Yes	☐ No
6. I can tell him what I usually read (kinds of books, newspapers, magazines, textbooks etc.)	☐ Yes	☐ No
7. I can ask him what newspapers there are in his own country.	☐ Yes	☐ No
8. I can tell him what I do in my free time.	☐ Yes	☐ No
9. I can ask him how to get to a certain place by public transport.	☐ Yes	☐ No
10. I can tell him what I think of art galleries.	☐ Yes	☐ No
11. I can ask him about the price of a ticket for a certain football match.	☐ Yes	☐ No
12. I can tell him about things that might interest a tourist in my home region.	☐ Yes	☐ No
13. I can ask him questions about traffic-rules in his own country.	☐ Yes	☐ No
14. I can say something about social security in my country (old-age pensions, medical care, etc.)	☐ Yes	☐ No
15. I can tell him what sort of government we have in my country.	☐ Yes	☐ No

16. I can say something about my political views and tell him whether I support a political party. ☐ Yes ☐ No

17. I can tell him how I feel at the moment (if I am hungry, tired, ill, etc.) ☐ Yes ☐ No

18. I can ask him to help me arrange an appointment with a doctor. ☐ Yes ☐ No

19. I can tell him that I take medicine regularly. ☐ Yes ☐ No

20. I can tell him that I am tired and need some rest. ☐ Yes ☐ No

21. I can ask him to repeat slowly what he just said. ☐ Yes ☐ No

22. I can ask him about the pronunciation of a certain word. ☐ Yes ☐ No

23. I can ask him to characterise the climate in his country. ☐ Yes ☐ No

24. I can ask him if he knows the approximate price of a certain piece of clothing in his own country. ☐ Yes ☐ No

25. I can inform him about where he can have his car serviced. ☐ Yes ☐ No

26. I can ask him to ring me up some time. ☐ Yes ☐ No

27. I can ask for his telephone number and give my own number. ☐ Yes ☐ No

28. I can tell him where he can change foreign money. ☐ Yes ☐ No

29. I can describe weather-conditions in the four seasons in my own country. ☐ Yes ☐ No

30. I can tell him where he can eat and drink. ☐ Yes ☐ No

Total number of crosses: ___ Yes ___ No

Evaluation: If your total number of Yes crosses is [25] or above, and if your judgement of your own language ability is fairly accurate, you are likely to have reached Threshold Level in English.

APPENDIX E

SELF-ASSESSMENT FORM (Type E)

SELF-ASSESSMENT FORM (Type E)

Language Functions and Exponents

This sample questionnaire is devised on the basis of the specification of language functions found in the English T-level document (van Ek, 1975). The functions are illustrated by exponents, also taken from the T-level document. Only exponents marked for productive use (P) are included. It is to be noted that this list only represents a random sample of the total inventory set out in the document.

Self-management by this method of course needs to be done on the basis of a native-language questionnaire. The form presented here is only to be seen as a possible model, not as a final product to be used as it stands. If used as a model it is to be translated into the learner's mother tongue.

SELF-ASSESSMENT FORM (Type E)

(To be translated into the learner's native language)

Instructions: Below you will find a number of statements concerning things that one can express in the language. Each statement is followed by two examples. Indicate your estimated command of the language by putting a cross in the appropriate box (YES or NO). Note that the sentences after the statements are examples only. If you can say the same thing in a similar way, you should put the cross in the YES box.

1.	I can ask a person if he or she is of the same opinion as I am. Do you agree? Don't you think so?	☐ Yes	☐ No
2.	I can offer to do something. Can I help you? Shall I shut the window?	☐ Yes	☐ No
3.	I can give a person permission to do something. You may smoke here. That's all right.	☐ Yes	☐ No
4.	I can ask others whether they have permission to do something. Are you allowed to walk here? Are you allowed to smoke in this room?	☐ Yes	☐ No
5.	I can decline an invitation. I'm afraid I cannot come. No, thank you.	☐ Yes	☐ No
6.	I can express pleasure. This is very nice. I like coffee very much.	☐ Yes	☐ No
7.	I can express satisfaction. This is very good. This is just what I want.	☐ Yes	☐ No
8.	I can express disappointment. That's a great pity. I am very sorry to hear that.	☐ Yes	☐ No
9.	I can express preference. I prefer sleeping. I'd rather not tell you.	☐ Yes	☐ No

10.	I can express sympathy. I am sorry to hear that. I am very glad to hear that.	☐ Yes	☐ No
11.	I can ask about somebody's intention. Are you going to stay long? Do you intend to go back early?	☐ Yes	☐ No
12.	I can apologize. I am very sorry. Excuse me, please.	☐ Yes	☐ No
13.	I can express disapproval. You shouldn't laugh. It's not very nice.	☐ Yes	☐ No
14.	I can ask others to do something. Would you please give me a hand? Can I have this shirt washed, please?	☐ Yes	☐ No
15.	I can advise others to do something. You should try this one. Why don't you join us?	☐ Yes	☐ No
16.	I can warn others to take care. Look out! Be careful!	☐ Yes	☐ No
17.	I can instruct others to do something. You turn left when you get to the church. Don't shut this door.	☐ Yes	☐ No
18.	I can take leave of somebody. I'll see you tomorrow. Good night.	☐ Yes	☐ No
19.	I can propose a toast. Here's to Charlie! Cheers!	☐ Yes	☐ No
20.	I can express approval. Good! That's fine!	☐ Yes	☐ No

Total number of crosses: ___ Yes ___ No

Evaluation: If your total number of Yes crosses is above 15, and if your judgement of your own language ability is fairly accurate, you are likely to have reached Threshold Level in English.